A NOTEBOOK ON VARIANCE COMPONENTS:

A DETAILED DESCRIPTION OF RECENT METHODS OF ESTIMATING VARIANCE COMPONENTS, WITH APPLICATIONS IN ANIMAL BREEDING

BU-640-M

SECOND DRAFT
February, 1978

S. R. Searle, Biometrics Unit
and
R. L. Quaas, Animal Science Department
Cornell University, Ithaca, New York

Abstract

The statistical methodology of variance components estimation has long enjoyed uses in many fields of application, especially in animal breeding and population genetics. For more than twenty years prior to 1967, the specific methods available were all based on the same theme, but in the succeeding decade several new methods have been developed that depart quite radically from this theme. They involve a considerable corpus of algebra and underlying mathematics. This Notebook describes these methods with all attendant details of the algebra, alternative forms of the results and relationships between them, as well as peripheral topics that help in understanding and/or computing the estimators. Computing difficulties are alluded to, but not considered at any length. The aim of the Notebook is to give a detailed description of the ML, REML, MINQUE, I-MINQUE and MIVQUE methods of estimating variance components in a unifying notation, to show the underlying mathematics and to illustrate the methods in the context of animal breeding applications.

Paper No. BU-640-M in the Biometrics Unit Mimeo Series.

TABLE OF CONTENTS

Page

Page

Page

0. INTRODUCTION

Variance components estimation as a statistical method was, for several decades, the poor step-child of analysis of variance, but in recent years the subject has generated widespread interest. Until 1967, methodology was based on equating sums of squares to their expected values, as proposed in the first papers on variance components, by Daniels [1939] and Winsor and Clarke [1940]. For what is nowadays usually called balanced data (having equal numbers of observations in the subclasses), this method involves sums of squares associated with traditional analysis of variance. Its use in a wide variety of models was first promulgated by the Anderson and Bancroft [1952] book, and useful minimum variance properties were derived in the late 1950's and early 1960's by Graybill and co-workers (e.g., Graybill and Wortham [1956] and Graybill and Hultquist [1961].

Data having unequal numbers of observations in the subclasses, possibly with some (or maybe many) subclasses with no data at all, are called unbalanced data. Estimation of variance components from such data has proven to be far more difficult than from balanced data. Henderson [1953] is a landmark paper for this situation, with its three methods of estimation based on the same principle as is used for balanced data, generalized to the extent of equating quadratic forms of the observations (rather than just sums of squares) to their expected values. This yields unbiased estimators — which are available by this means in embarrassing abundance. But they have few, if any, other attractive properties — save, in some cases, of being relatively easy to compute. The years following Henderson [1953] saw considerable use, expansion and explanation of his methods, together with exploration of their properties (see, e.g., Searle [1971a, b]), but there were no really new developments until Hartley and Rao [1967] described maximum likelihood procedures — based, as is often the case, on normality assumptions. Since then there has been a host of new methods, not only

ML: maximum likelihood,

but also

REML: restricted maximum likelihood

MINQUE: minimum norm quadratic unbiased estimation

I-MINQUE: iterative MINQUE

MIVQUE: minimum variance quadratic unbiased estimation

and, no doubt, other alphabetic horrors. In addition there are tangential topics such as the MME's (mixed model equations) and the dispersion-mean model.

The foundation for development of many of these things is a large corpus of matrix algebra. This is complicated in the literature by numerous notations that look sufficiently alike to add the traditional amount of confusion; and, hanging like a thunder cloud over everything, are numerical and computing problems involved with very large data sets, large and sparse matrices, and the solving of non-linear equations subject to non-linear (non-negativity) constraints. This Notebook directs its attention not to the last of these things (computing) but to the first two: it aims at describing development of the methods in a unifying notation, along with all necessary algebraic details to support that development, including alternative, but equivalent expressions for and also relationships between, the methods.

There is growing need to have this algebraic development readily available, despite the mathematical tedium of it when presented <u>in extenso</u>. It is useful for teaching, for developing new computing algorithms and understanding old ones; and, especially to users of variance components, for their being able to properly understand the estimation methods currently available. Many users have adequate ability for this, but maybe insufficient time for "filling in the gaps" that exist in published research papers which, in order to save publication space, have had to resort to phrases like "it can be shown that ···" and "after lengthy algebraic manipulation we find that ···". Making good on these phrases is one aim of this Notebook.

Users are also faced with such an apparent plethora of estimation methods that the question "what method should be used?" has become even more difficult to answer than some years ago, when just the three Henderson methods were available. A minimum first step towards answering it is to know what the available methods are. Also, since there are relationships between some of the methods, users feel some dissatisfaction at reading of these relationships without having their detailed description; e.g., "The first round of REML iteration yields a MINQUE estimate." This Notebook therefore aims at describing the methods in detail and showing their relationships.

This is not a review paper. Extensive reviews are to be found in Searle [1971a] and Harville [1975, 1977]. Considerable detail of the Henderson methods is in Searle [1971b], so that attention is confined to methods that have been proposed since then, namely ML, REML, MINQUE, I-MINQUE and MIVQUE. Particular, but not exclusive, attention is also paid to the development of specific results given in Harville [1977], which is a substantive review of the topics dealt with here in detail.

The Notebook is in two major parts. Part I deals with development of the methods. Part II illustrates some of their uses and applications to animal breeding.

Chapter 1

MODELS AND NOTATION

Variance components estimation is based upon the statistical linear model. In its most general form, a popular representation of this is

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{e} \, , \tag{1.1}$$

where $\underset{\sim}{y}$ is a vector of observations, $\underset{\sim}{\beta}$ is a vector of unknown parameters, $\underset{\sim}{X}$ is a known matrix and $\underset{\sim}{e}$ is a vector of unobservable random errors. $\underset{\sim}{X}$ is often a design, or incidence, matrix of 0's and 1's, but it can also include columns of covariates.

The usual form of the linear model is (1.1) when $\underset{\sim}{\beta}$ is a vector of fixed effects. But when $\underset{\sim}{\beta}$ includes both fixed and random effects, to be denoted hereafter by $\underset{\sim}{\alpha}$ and $\underset{\sim}{b}$ respectively, we rewrite the model as in equation (2.1) of Harville [1977] (which we subsequently refer to as [H2.1], for example). The model is then

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\alpha} + \underset{\sim}{Z}\underset{\sim}{b} + \underset{\sim}{e} \tag{1.2}$$

with the following definitions:

$\underset{\sim}{y}_{N\times 1}$ is a vector of N observations,

$\underset{\sim}{\alpha}_{p\times 1}$ is a vector of p fixed effect parameters,

$\underset{\sim}{X}_{N\times p}$ is a known matrix, of rank $p^* \leq p < N$,

$\underset{\sim}{b}_{q\times 1}$ is a vector of q random effects,

$\underset{\sim}{Z}_{N\times q}$ is a known matrix, often consisting of just 0's and 1's

and

$\underset{\sim}{e}_{N\times 1}$ is a vector of residual error terms.

In order to identify the variance components corresponding to the random effects in $\underset{\sim}{b}$, this vector is partitioned as

$$\underset{\sim}{b}' = [\underset{\sim}{b}_1' \;\; \underset{\sim}{b}_2' \cdots \underset{\sim}{b}_i' \cdots \underset{\sim}{b}_c'] = \{\underset{\sim}{b}_i'\} \quad \text{for } i = 1, \cdots, c \tag{1.3}$$

where

$\underset{\sim}{b}_i$ is a vector of the q_i effects of the i'th random factor.

In this context a factor means either a main effects factor or an interaction factor. (For example, in the 2-way crossed classification there are 3 factors, rows, columns and interaction.) Corresponding to $\underset{\sim}{b}_i$ of (1.3) is the incidence matrix $\underset{\sim}{Z}_i$ of order $N \times q_i$, so that in (1.2)

$$\underset{\sim}{Z} = [\underset{\sim}{Z}_1 \;\; \underset{\sim}{Z}_2 \cdots \underset{\sim}{Z}_i \cdots \underset{\sim}{Z}_c] = \{\underset{\sim}{Z}_i\} \quad \text{for } i = 1, \cdots, c \tag{1.4}$$

thus giving

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\alpha} + \sum_{i=1}^{c} \underset{\sim}{Z}_i\underset{\sim}{b}_i + \underset{\sim}{e} \, . \tag{1.5}$$

Distributional properties imputed to (1.5) are as follows:

$$E(\underset{\sim}{b}_i) = 0, \quad E(\underset{\sim}{e}) = 0, \quad E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{\alpha} \tag{1.6}$$

where E represents expectation; and

$$\mathrm{var}(\underset{\sim}{b}_i) = \sigma_i^2 \underset{\sim}{I}_{q_i} \qquad \mathrm{cov}(\underset{\sim}{b}_i, \underset{\sim}{b}_j') = \underset{\sim}{0} \quad i \neq j \tag{1.7}$$

$$\mathrm{var}(\underset{\sim}{e}) = \underset{\sim}{R} \qquad \mathrm{cov}(\underset{\sim}{b}_i, \underset{\sim}{e}') = \underset{\sim}{0} \, , \tag{1.8}$$

where $\mathrm{var}(\underset{\sim}{e})$, for example, is the variance-covariance matrix of $\underset{\sim}{e}$, and $\mathrm{cov}(\underset{\sim}{b}_i, \underset{\sim}{e}')$ is the matrix of covariances of the elements of $\underset{\sim}{b}_i$ with those of $\underset{\sim}{e}'$. From (1.7), the variance of $\underset{\sim}{b}$ in (1.3) is

$$\mathrm{var}(\underset{\sim}{b}) = \underset{\sim}{D} = \begin{bmatrix} \sigma_1^2 \underset{\sim}{I}_{q_1} & & \underset{\sim}{0} \\ & \ddots & \\ \underset{\sim}{0} & & \sigma_c^2 \underset{\sim}{I}_{q_c} \end{bmatrix} \tag{1.9}$$

which can also be written as

$$\underset{\sim}{D} \equiv \sum_{i=1}^{c}{}^{+} \sigma_i^2 \underset{\sim}{I}_{q_i} \tag{1.10}$$

$$\equiv \text{diag}\{\sigma_i^2 \underset{\sim}{I}_{q_i}\} \quad \text{for } i = 1, \cdots, c \tag{1.11}$$

where, in (1.10), Σ^+ represents a direct sum and, in (1.11), diag{ } represents a diagonal matrix, in this case of sub-matrices $\sigma_i^2 \underset{\sim}{I}_{q_i}$. Then (1.8) and (1.9) applied to (1.5) give the variance of $\underset{\sim}{y}$, defined as $\underset{\sim}{V}$, to be

$$\text{var}(\underset{\sim}{y}) \equiv \underset{\sim}{V} = E(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})' = \underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{Z}' + \underset{\sim}{R} . \tag{1.12}$$

By the nature of $\underset{\sim}{D}$ in (1.10), and on using (1.4), this is also

$$\underset{\sim}{V} = \sum_{i=1}^{c} \sigma_i^2 \underset{\sim}{Z}_i \underset{\sim}{Z}_i' + \underset{\sim}{R} . \tag{1.13}$$

Sometimes the definition

$$\underset{\sim}{V}_i = \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \tag{1.14}$$

is also used, giving

$$\underset{\sim}{V} = \sum_{i=1}^{c} \sigma_i^2 \underset{\sim}{V}_i + \underset{\sim}{R} . \tag{1.15}$$

Note that $\underset{\sim}{V}$ is symmetric, of order N, and positive definite, so that

$$\underset{\sim}{V}' = \underset{\sim}{V} \quad \text{and} \quad \underset{\sim}{V}^{-1} = \underset{\sim}{L}'\underset{\sim}{L} \text{ for some non-singular } \underset{\sim}{L} . \tag{1.16}$$

It is common practice to take

$$\text{var}(\underset{\sim}{e}) = \underset{\sim}{R} = \sigma_0^2 \underset{\sim}{I}_N \tag{1.17}$$

where σ_0^2 is then the variance of each element of $\underset{\sim}{e}$. Alternative symbols often used for σ_0^2 are σ_e^2 and σ^2; but σ_0^2 shall be used here. Then (1.13) and (1.15) are

$$\underset{\sim}{V} = \sum_{i=1}^{c} \sigma_i^2 \underset{\sim}{Z}_i \underset{\sim}{Z}_i' + \sigma_0^2 \underset{\sim}{I}_N \tag{1.18}$$

$$= \sum_{i=1}^{c} \sigma_i^2 \underset{\sim}{V}_i + \sigma_0^2 \underset{\sim}{I}_N \, . \tag{1.19}$$

The definition

$$\gamma_i = \sigma_i^2/\sigma_0^2 \quad \text{for } i = 1, \cdots, c \tag{1.20}$$

is also used, along with defining $\underset{\sim}{H}$ by

$$\underset{\sim}{V} = \sigma_0^2 \underset{\sim}{H} \text{ , i.e., } \underset{\sim}{H} = \sigma_0^{-2} \underset{\sim}{V} \tag{1.21}$$

so that from (1.18)

$$\underset{\sim}{H} = \sum_{i=1}^{c} \gamma_i \underset{\sim}{Z}_i \underset{\sim}{Z}_i' + \underset{\sim}{I}_N \, . \tag{1.22}$$

A compact way of writing the model (1.5) is to define $\underset{\sim}{e}$ as another $\underset{\sim}{b}_i$, namely $\underset{\sim}{b}_0$, and the corresponding $\underset{\sim}{Z}_0$ as $\underset{\sim}{I}_N$. Then, on using

$$\underset{\sim}{b}_0 = \underset{\sim}{e}, \quad q_0 = N \quad \text{and} \quad \underset{\sim}{Z}_0 = \underset{\sim}{I}_N \tag{1.23}$$

we have (1.5) and (1.17) as

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\alpha} + \sum_{i=0}^{c} \underset{\sim}{Z}_i \underset{\sim}{b}_i \tag{1.24}$$

and

$$\underset{\sim}{V} = \sum_{i=0}^{c} \sigma_i^2 \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \, . \tag{1.25}$$

It is sometimes convenient to define

$$\gamma_0 \equiv \sigma_0^2 \, . \tag{1.26}$$

Vector notation also used for the σ^2's and γ's is then

$$\underset{\sim}{\sigma}^2 = \begin{bmatrix} \sigma_1^2 \\ \vdots \\ \sigma_c^2 \end{bmatrix}, \quad \dot{\underset{\sim}{\sigma}}^2 = \begin{bmatrix} \sigma_0^2 \\ \sigma_1^2 \\ \vdots \\ \sigma_c^2 \end{bmatrix}, \quad \underset{\sim}{\gamma} = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_c \end{bmatrix} \text{ and } \dot{\underset{\sim}{\gamma}} = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_c \end{bmatrix} \text{ with } \gamma_0 \equiv \sigma_0^2. \tag{1.27}$$

In the same way that $\dot{\underset{\sim}{\sigma}}^2$ is an extension of $\underset{\sim}{\sigma}^2$, we also define $\dot{\underset{\sim}{b}}$, $\dot{\underset{\sim}{Z}}$ and $\dot{\underset{\sim}{D}}$:

$$\dot{\underset{\sim}{b}} = \begin{bmatrix} \underset{\sim}{b}_0 \\ \underset{\sim}{b}_1 \\ \vdots \\ b_c \end{bmatrix}, \quad \dot{\underset{\sim}{Z}}' = \begin{bmatrix} \underset{\sim}{Z}_0' \\ \underset{\sim}{Z}_1' \\ \vdots \\ \underset{\sim}{Z}_c' \end{bmatrix} \quad \text{and} \quad \dot{\underset{\sim}{D}} = \begin{bmatrix} \sigma_0^2 \underset{\sim}{I}_{q_0} & & & \\ & \sigma_1^2 \underset{\sim}{I}_{q_1} & & \\ & & \ddots & \\ & & & \sigma_c^2 \underset{\sim}{I}_{q_c} \end{bmatrix} \tag{1.28}$$

and so write the model as

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\alpha} + \dot{\underset{\sim}{Z}}\dot{\underset{\sim}{b}} \tag{1.29}$$

with

$$\underset{\sim}{V} = \dot{\underset{\sim}{Z}}\dot{\underset{\sim}{D}}\dot{\underset{\sim}{Z}}' . \tag{1.30}$$

A variation of the subscript 0 is to use $c+1$ in its place; e.g., $\sigma_{c+1}^2 \equiv \sigma_0^2 \equiv \sigma_e^2$, as is done in Harville [1975, 1977]. We find 0 to be more compact.

Chapter 2

RESULTS IN MATRIX ALGEBRA

Matrix algebra plays an important role in describing the newer methods of estimating variance components, and especially in the derivation of these methods. Presentation of required results could be achieved either by: (i) introducing them just as needed, throughout the mainstream of the Notebook or (ii) by collecting them in one place, either in an appendix or an early chapter. These two styles complement one another: (i) has the advantage that presentation of each matrix result is motivated by context, but a consequence is that the results are widely scattered, with little cohesion, and out of context are difficult to find and refer to. In contrast, in (ii), the results are presented without motivation but are easy to find and to use. We have chosen to collect results in one place, preferring an early chapter to an appendix on grounds of logical development. Readers interested primarily in variance components, and who are prepared to take matrix results on faith, may therefore treat this chapter as an appendix and go straight to Chapter 3.

2.1. DIFFERENTIATION

Differentiation of matrices and quadratic forms is used in all of the ML, REML and MINQUE methods of estimating variance components. Basic results are given here, and more specialized results in Sections 2.7 and 2.8. For a vector of distinct variables $\underset{\sim}{x}$,

$$\frac{\partial(\underset{\sim}{A}\underset{\sim}{x})}{\partial\underset{\sim}{x}} = \underset{\sim}{A}' \quad \text{and, for symmetric } \underset{\sim}{A}, \quad \frac{\partial(\underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x})}{\partial\underset{\sim}{x}} = 2\underset{\sim}{A}\underset{\sim}{x}\ , \tag{2.1}$$

(Searle [1966, pp. 204-206]). Also, when elements of $\underset{\sim}{A}$ are functions of the scalar u,

$$\frac{\partial \underset{\sim}{A}^{-1}}{\partial u} = -\underset{\sim}{A}^{-1} \frac{\partial \underset{\sim}{A}}{\partial u} \underset{\sim}{A}^{-1} , \tag{2.2}$$

is a standard result. And, very simply, for $\underset{\sim}{V}$ of (1.24)

$$\frac{\partial \underset{\sim}{V}}{\partial \sigma_i^2} = \underset{\sim}{Z}_i \underset{\sim}{Z}_i' ; \tag{2.3}$$

and

$$\frac{\partial}{\partial \sigma_i^2} \log|\underset{\sim}{V}| = \mathrm{tr}\left(\underset{\sim}{V}^{-1} \frac{\partial \underset{\sim}{V}}{\partial \sigma_i^2}\right) = \mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') \tag{2.4}$$

as in Searle [1970, p. 510].

2.2. DETERMINANTS

Determinants occur in likelihood functions that are the basis for ML and REML. A useful first result is

$$\begin{vmatrix} \underset{\sim}{A} & \underset{\sim}{B} \\ \underset{\sim}{C} & \underset{\sim}{D} \end{vmatrix} = |\underset{\sim}{A}|\ |\underset{\sim}{D} - \underset{\sim}{C}\underset{\sim}{A}^{-1}\underset{\sim}{B}| = |\underset{\sim}{D}|\ |\underset{\sim}{A} - \underset{\sim}{B}\underset{\sim}{D}^{-1}\underset{\sim}{C}| . \tag{2.5}$$

The first equality in (2.5) holds for $\underset{\sim}{A}$ non-singular and the second for $\underset{\sim}{D}$ non-singular (Searle [1966, p. 96]). Another useful result is

$$|\underset{\sim}{A}\quad \underset{\sim}{B}| = \begin{vmatrix} \underset{\sim}{A}' \\ \underset{\sim}{B}' \end{vmatrix} = \left|\begin{pmatrix} \underset{\sim}{A}' \\ \underset{\sim}{B}' \end{pmatrix} (\underset{\sim}{A}\quad \underset{\sim}{B})\right|^{\frac{1}{2}} = \begin{vmatrix} \underset{\sim}{A}'\underset{\sim}{A} & \underset{\sim}{A}'\underset{\sim}{B} \\ \underset{\sim}{B}'\underset{\sim}{A} & \underset{\sim}{B}'\underset{\sim}{B} \end{vmatrix}^{\frac{1}{2}} . \tag{2.6}$$

When (2.6) is non-zero, $\underset{\sim}{A}$ and $\underset{\sim}{B}$ each have full column rank and so $\underset{\sim}{A}'\underset{\sim}{A}$ and $\underset{\sim}{B}'\underset{\sim}{B}$

are positive definite. Then (2.5) can be applied to (2.6) giving

$$|\mathbf{A}\ \ \mathbf{B}|^2 = |\mathbf{A}'\mathbf{A}|\ |\mathbf{B}'\mathbf{B} - \mathbf{B}'\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{B}| = |\mathbf{B}'\mathbf{B}|\ |\mathbf{A}'\mathbf{A} - \mathbf{A}'\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{A}| \ .$$

Also useful is

$$|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}| \ , \tag{2.7}$$

established by applying (2.5) to the determinant of

$$\begin{bmatrix} \mathbf{I} & -\mathbf{A} \\ \mathbf{B} & \mathbf{I} \end{bmatrix} .$$

2.3. GENERALIZED INVERSES

Generalized inverse matrices arise in the solution of least squares equations. A generalized inverse of $\mathbf{X}'\mathbf{X}$ is $(\mathbf{X}'\mathbf{X})^-$ defined by

$$\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X} = \mathbf{X}'\mathbf{X} \ . \tag{2.8}$$

Useful properties include the following:

(i) $(\mathbf{X}'\mathbf{X})^{-\prime}$ is also a generalized inverse of $\mathbf{X}'\mathbf{X}$.

(ii) $\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ is invariant to the choice of $(\mathbf{X}'\mathbf{X})^-$.

(iii) $\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ is symmetric. (2.9)

(iv) $\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X} = \mathbf{X}$.

(v) $r[(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X}] = r(\mathbf{X})$.

(See Searle [1971b, pp. 12 and 20].)

The Moore-Penrose form of generalized inverse of any non-null matrix $\mathbf{B}$ is $\mathbf{B}^+$ where

$$\mathbf{B}^+ \text{ is unique and } \begin{array}{l} \mathbf{BB}^+\mathbf{B} = \mathbf{B} \\ \mathbf{B}^+\mathbf{BB}^+ = \mathbf{B}^+ \end{array} \text{ and } \begin{array}{l} \mathbf{BB}^+ = (\mathbf{BB}^+)' \\ \mathbf{B}^+\mathbf{B} = (\mathbf{B}^+\mathbf{B})' \end{array} . \tag{2.10}$$

2.4. TRACE OPERATIONS

Trace operations occur in their application to quadratic forms; e.g.,

$$\mathrm{tr}(\underset{\sim}{ABC}) = \mathrm{tr}(\underset{\sim}{CAB}) \quad \text{applied to} \quad \underset{\sim}{x}'\underset{\sim}{Ay} = \mathrm{tr}(\underset{\sim}{x}'\underset{\sim}{Ay}), = \mathrm{tr}(\underset{\sim}{Ayx}') . \tag{2.11}$$

These results receive repeated use, especially the latter in the form

$$E(\underset{\sim}{y}'\underset{\sim}{Ay}) = E\ \mathrm{tr}(\underset{\sim}{Ayy}') = \mathrm{tr}\ \underset{\sim}{A}\ E(\underset{\sim}{yy}') = \mathrm{tr}(\underset{\sim}{AV}) + \underset{\sim}{\mu}'\underset{\sim}{A\mu} \tag{2.12}$$

where E is an expectation operator, and $\underset{\sim}{y}$ has mean $\underset{\sim}{\mu}$ and variance-covariance matrix $\underset{\sim}{V}$, i.e., $\underset{\sim}{y}$ is distributed $(\underset{\sim}{\mu}, \underset{\sim}{V})$, or $\underset{\sim}{y} \sim (\underset{\sim}{\mu}, \underset{\sim}{V})$, but not necessarily normally distributed.

2.5. THE INVERSE OF $\underset{\sim}{V}$

The inverse of $\underset{\sim}{V} = \mathrm{var}(\underset{\sim}{y})$ arises in likelihood functions and in generalized least squares equations. Because, as in (1.12), $\underset{\sim}{V}$ has the form $\underset{\sim}{V} = \underset{\sim}{ZDZ}' + \underset{\sim}{R}$, there are various expressions for its inverse, one being

$$\underset{\sim}{V}^{-1} = (\underset{\sim}{ZDZ}' + \underset{\sim}{R})^{-1} = \underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{Z}(\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1} \tag{2.13}$$

when $\underset{\sim}{D}$ and $\underset{\sim}{R}$ are non-singular. Verification comes from multiplying the right-hand side of (2.13) by $\underset{\sim}{V}$ and obtaining $\underset{\sim}{I}$:

$$\begin{aligned}
&\underset{\sim}{V}[\underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{Z}(\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}] \\
&\quad = (\underset{\sim}{ZDZ}' + \underset{\sim}{R})[\underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{Z}(\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1}] \\
&\quad = \underset{\sim}{I} + \underset{\sim}{ZDZ}'\underset{\sim}{R}^{-1} - (\underset{\sim}{ZDZ}'\underset{\sim}{R}^{-1}\underset{\sim}{Z} + \underset{\sim}{Z})(\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1} \\
&\quad = \underset{\sim}{I} + \underset{\sim}{ZDZ}'\underset{\sim}{R}^{-1} - \underset{\sim}{ZD}(\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z} + \underset{\sim}{D}^{-1})(\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1} \\
&\quad = \underset{\sim}{I} .
\end{aligned}$$

An alternative form to (2.13) which, as Harville [1977] points out, does not require $\underset{\sim}{D}$ to be non-singular — which (2.13) does require — is

$$\underset{\sim}{V}^{-1} = \underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1} . \tag{2.14}$$

Although this form does not appear to be symmetric, it is, of course, because $\underset{\sim}{V}$ is symmetric. Verification of (2.14) is available in the same manner as is that of (2.13).

Note from (2.14) that

$$\begin{aligned} \underset{\sim}{Z}'\underset{\sim}{V}^{-1} &= [\underset{\sim}{I} - \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D})^{-1}]\underset{\sim}{Z}'\underset{\sim}{R}^{-1} \\ &= [\underset{\sim}{I} - (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D} - \underset{\sim}{I})(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D})^{-1}]\underset{\sim}{Z}'\underset{\sim}{R}^{-1} \\ &= (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1} , \end{aligned} \tag{2.15}$$

which is the result below [H3.6].

The non-singularity of $\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}$ in (2.14) is assured by the following argument. $\underset{\sim}{D}$ of (1.9) is a dispersion (variance-covariance) matrix and so is positive semi-definite with $\underset{\sim}{D} = \underset{\sim}{B}\underset{\sim}{B}'$ for some $\underset{\sim}{B}$ of full column rank; similarly, $\underset{\sim}{R}$ is positive definite with $\underset{\sim}{R}^{-1} = \underset{\sim}{G}'\underset{\sim}{G}$ for some $\underset{\sim}{G}$. Hence, with all matrices being real, $|\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}| = |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{G}'\underset{\sim}{G}\underset{\sim}{Z}\underset{\sim}{B}\underset{\sim}{B}'| = |\underset{\sim}{I} + \underset{\sim}{B}'\underset{\sim}{Z}'\underset{\sim}{G}'\underset{\sim}{G}\underset{\sim}{Z}\underset{\sim}{B}|$, by (2.7), $= |\underset{\sim}{I} + \underset{\sim}{K}\underset{\sim}{K}'|$ for $\underset{\sim}{K}' = \underset{\sim}{G}\underset{\sim}{Z}\underset{\sim}{B}$, and, by Lemma 8 of Searle [1971b, p. 24], the matrix $\underset{\sim}{I} + \underset{\sim}{K}\underset{\sim}{K}'$ has full rank. Hence $|\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}| = |\underset{\sim}{I} + \underset{\sim}{K}\underset{\sim}{K}'| \neq 0$, and so, defining $\underset{\sim}{T}^*$,

$$\underset{\sim}{T}^* = (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D})^{-1} \text{ exists.} \tag{2.16}$$

This development is due to Harville [1976] with assistance from Anderson [1977].

2.6. SOME SPECIAL MATRICES

Certain matrices that play prominent roles in what follows are now defined and some of their many properties presented.

a. The Matrices $\underset{\sim}{M}$, $\underset{\sim}{P}$ and $\underset{\sim}{S}$

The well-known matrix

$$\underset{\sim}{M} \equiv \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}' \tag{2.17}$$

is important in the REML and MINQUE methodology. It is, of course, symmetric and idempotent:

$$\underset{\sim}{M} = \underset{\sim}{M}' = \underset{\sim}{M}^2 \tag{2.18}$$

$$\underset{\sim}{M}\underset{\sim}{X} = \underset{\sim}{0} \quad \text{and} \quad \underset{\sim}{X}'\underset{\sim}{M} = \underset{\sim}{0}\ , \tag{2.19}$$

and

$$\text{tr}(\underset{\sim}{M}) = r(\underset{\sim}{M}) = N - p^* \ , \tag{2.20}$$

where $r(\underset{\sim}{M})$ represents the rank of $\underset{\sim}{M}$. Because, by (2.18) and (2.20), $\underset{\sim}{M}$ is symmetric and idempotent of rank $N - p^*$, its canonical form under orthogonal similarity is

$$\underset{\sim}{U}\underset{\sim}{M}\underset{\sim}{U}' = \begin{bmatrix} \underset{\sim}{I}_{N-p^*} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} \tag{2.21}$$

where $\underset{\sim}{U}$ is orthogonal. Letting $\underset{\sim}{A}$ be the first $N - p^*$ columns of $\underset{\sim}{U}'$, it is then not difficult to show that

$$\underset{\sim}{M} = \underset{\sim}{A}\underset{\sim}{A}' \quad \text{and} \quad \underset{\sim}{A}'\underset{\sim}{A} = \underset{\sim}{I}\ . \tag{2.22}$$

Post-multiplying $\underset{\sim}{M} = \underset{\sim}{A}\underset{\sim}{A}'$ by $\underset{\sim}{A}$ and using $\underset{\sim}{A}'\underset{\sim}{A} = \underset{\sim}{I}$ gives

$$\underset{\sim}{M}\underset{\sim}{A} = \underset{\sim}{A} \quad \text{and} \quad \underset{\sim}{A}'\underset{\sim}{M} = \underset{\sim}{A}'\ ; \tag{2.23}$$

and then post-multiplying $\underset{\sim}{A}'\underset{\sim}{M} = \underset{\sim}{A}'$ by $\underset{\sim}{X}$ and using $\underset{\sim}{M}\underset{\sim}{X} = \underset{\sim}{0}$ of (2.19) gives $\underset{\sim}{A}'\underset{\sim}{X} = \underset{\sim}{0}$.

Furthermore, because $\underset{\sim}{U}'$ is orthogonal and hence non-singular, $\underset{\sim}{A}'$ has full row rank. Hence, in addition to (2.22) we also have

$$(\underset{\sim}{A}')_{N-p^{*}\times N} \text{ of full row rank } N - p^{*}, \text{ with } \underset{\sim}{A}'\underset{\sim}{X} = \underset{\sim}{0} \,. \tag{2.24}$$

With the aid of (2.23) it also is easy to verify that $\underset{\sim}{A}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{A}$ is the Moore-Penrose inverse, see (2.10), of $\underset{\sim}{M}\underset{\sim}{V}\underset{\sim}{M}$:

$$(\underset{\sim}{M}\underset{\sim}{V}\underset{\sim}{M})^{+} = \underset{\sim}{A}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{A}' \,. \tag{2.25}$$

The inverse, $(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}$ exists because $\underset{\sim}{A}'$ has full row rank and $\underset{\sim}{V}$ is positive definite.

A generalization of $\underset{\sim}{M}$ is

$$\underset{\sim}{P} = \underset{\sim}{V}^{-1} - \underset{\sim}{V}^{-1}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1} \,, \tag{2.26}$$

since when $\underset{\sim}{V} = \underset{\sim}{I}$, $\underset{\sim}{P}$ becomes $\underset{\sim}{M}$. MINQUE depends upon $\underset{\sim}{P}$, which has many properties, the simplest of which is, using (1.16) and (2.9),

$$\underset{\sim}{P}\underset{\sim}{X} = \underset{\sim}{0} \quad \text{and} \quad \underset{\sim}{X}'\underset{\sim}{P} = \underset{\sim}{0} \,. \tag{2.27}$$

Clearly, by multiplying (2.26) by $\underset{\sim}{V}$

$$\underset{\sim}{P}\underset{\sim}{V} = \underset{\sim}{I} - \underset{\sim}{V}^{-1}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}' \tag{2.28}$$

so that on using (2.27)

$$\underset{\sim}{P}\underset{\sim}{V}\underset{\sim}{P} = \underset{\sim}{P} \tag{2.29}$$

and hence

$$(\underset{\sim}{P}\underset{\sim}{V})^{2} = \underset{\sim}{P}\underset{\sim}{V} \,. \tag{2.30}$$

Therefore

$$\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{V}) = r(\underset{\sim}{P}\underset{\sim}{V}) = r(\underset{\sim}{P}) = N - p^{*} \,. \tag{2.31}$$

Relationships between $\mathbf{M}$ and $\mathbf{P}$ start with

$$\mathbf{PM} = \mathbf{P} = \mathbf{MP} \tag{2.32}$$

based on (2.17) and (2.27). Post multiplying (2.28) by $\mathbf{M}$ and using (2.19) gives $\mathbf{PVM} = \mathbf{M}$, so that on post-multiplying (2.32) by $\mathbf{VM}$ we also have

$$\mathbf{PMVM} = \mathbf{PVM} = \mathbf{MPVM} = \mathbf{M}\ . \tag{2.33}$$

This, along with (2.32), immediately leads to establishing $\mathbf{P}$ as the Moore-Penrose inverse of $\mathbf{MVM}$:

$$(\mathbf{MVM})^{+} = \mathbf{P}\ ; \tag{2.34}$$

and since $(\mathbf{MVM})^{+}$ is unique, (2.25) and (2.34) give

$$\mathbf{P} = \mathbf{A}(\mathbf{A}'\mathbf{VA})^{-1}\mathbf{A}' = (\mathbf{MVM})^{+}\ . \tag{2.35}$$

If in $\mathbf{V} = \mathbf{ZDZ}' + \mathbf{R}$ of (1.12) the matrix $\mathbf{Z}$ is null then $\mathbf{V}$ becomes $\mathbf{R}$ and $\mathbf{P}$ of (2.26) becomes

$$\mathbf{S} \equiv \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{R}^{-1}, \quad = \mathbf{P} \text{ with } \mathbf{Z} = \mathbf{0}\ . \tag{2.36}$$

Results corresponding to (2.27) through (2.35) are an immediate consequence:

$$\mathbf{SX} = \mathbf{0} \quad \text{and} \quad \mathbf{X}'\mathbf{S} = \mathbf{0} \tag{2.37}$$

$$\mathbf{SR} = \mathbf{I} - \mathbf{R}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})\mathbf{X}' \tag{2.38}$$

$$\mathbf{SRS} = \mathbf{S} \tag{2.39}$$

$$(\mathbf{SR})^{2} = \mathbf{SR} \tag{2.40}$$

$$\mathrm{tr}(\mathbf{SR}) = r(\mathbf{SR}) = r(\mathbf{S}) = N - p^{*} \tag{2.41}$$

$$\mathbf{SM} = \mathbf{S} = \mathbf{MS} \tag{2.42}$$

$$\mathbf{SMRM} = \mathbf{SRM} = \mathbf{MSRM} = \mathbf{M} \tag{2.43}$$

$$(\mathbf{MRM})^+ = \mathbf{S} \tag{2.44}$$

$$\mathbf{S} = \mathbf{A}(\mathbf{A}'\mathbf{RA})^{-1}\mathbf{A}' \ . \tag{2.45}$$

A functional relationship between $\mathbf{P}$ and $\mathbf{S}$,

$$\mathbf{P} = \mathbf{S} - \mathbf{SZD}(\mathbf{I} + \mathbf{Z}'\mathbf{SZD})^{-1}\mathbf{Z}'\mathbf{S} \ , \tag{2.46}$$

which precedes [H3.7], is derived as follows. From (2.35) and (1.12)

$$\mathbf{P} = \mathbf{A}(\mathbf{A}'\mathbf{VA})^{-1}\mathbf{A}' = \mathbf{A}(\mathbf{A}'\mathbf{ZDZ}'\mathbf{A} + \mathbf{A}'\mathbf{RA})^{-1}\mathbf{A}' \ ,$$

and if in this expression we write the inverse matrix in a form analogous to $\mathbf{V}^{-1}$ of (2.14), we have

$$\mathbf{P} = \mathbf{A}\left\{(\mathbf{A}'\mathbf{RA})^{-1} - (\mathbf{A}'\mathbf{RA})^{-1}\mathbf{A}'\mathbf{ZD}[\mathbf{I} + \mathbf{Z}'\mathbf{A}(\mathbf{A}'\mathbf{RA})^{-1}\mathbf{A}'\mathbf{ZD}]^{-1}\mathbf{Z}'\mathbf{A}(\mathbf{A}'\mathbf{RA})^{-1}\right\}\mathbf{A}'$$

which, on using (2.45), is (2.46). Analogous to (2.16) define $\mathbf{T}$ as

$$\mathbf{T} = (\mathbf{I} + \mathbf{Z}'\mathbf{SZD})^{-1} \ , \tag{2.47}$$

noting that it does exist, by arguments similar to those used for establishing (2.16). With (2.47), $\mathbf{P}$ of (2.46) is then

$$\mathbf{P} = \mathbf{S} - \mathbf{SZDTZ}'\mathbf{S} \ . \tag{2.48}$$

b. Simplified Forms

We have already noted in (2.36) that $\mathbf{Z} = \mathbf{0} \Rightarrow \mathbf{P} = \mathbf{S}$, as is additionally evident in (2.48). Further simplified forms can also be listed:

$$\mathbf{V} = \mathbf{I} \quad \Rightarrow \mathbf{P} = \mathbf{M} \tag{2.49}$$

$$\mathbf{R} = \mathbf{I} \quad \Rightarrow \mathbf{S} = \mathbf{M} \tag{2.50}$$

$$\mathbf{X} = \mathbf{0} \quad \Rightarrow \mathbf{P} = \mathbf{V}^{-1} \text{ and } \mathbf{S} = \mathbf{R}^{-1} \tag{2.51}$$

$$\mathbf{R} = \sigma_0^2\mathbf{I} \Rightarrow \mathbf{S} = \mathbf{M}/\sigma_0^2 \tag{2.52}$$

$$\Rightarrow \mathbf{P} = [\mathbf{M} - \mathbf{MZD}(\sigma_0^2\mathbf{I} + \mathbf{Z}'\mathbf{MZD})^{-1}\mathbf{Z}'\mathbf{M}]/\sigma_0^2 \tag{2.53}$$

$$= [\mathbf{M} - \mathbf{MZ}(\sigma_0^2\mathbf{D}^{-1} + \mathbf{Z}'\mathbf{MZ})^{-1}\mathbf{Z}'\mathbf{M}]/\sigma_0^2 \tag{2.54}$$

when $\mathbf{D}^{-1}$ exists. Alternative forms, on defining

$$\mathbf{D}_0 = \mathbf{D}/\sigma_0^2 = \text{diag}\{\gamma_i\mathbf{I}_{q_i}\} \text{ for } i = 1, \cdots, c \tag{2.55}$$

are

$$\mathbf{P} = [\mathbf{M} - \mathbf{MZD}_0(\mathbf{I} + \mathbf{Z}'\mathbf{MZD}_0)^{-1}\mathbf{Z}'\mathbf{M}]/\sigma_0^2 \tag{2.56}$$

$$= [\mathbf{M} - \mathbf{MZ}(\mathbf{D}_0^{-1} + \mathbf{Z}'\mathbf{MZ})^{-1}\mathbf{Z}'\mathbf{M}]\sigma_0^2 \,. \tag{2.57}$$

c. Using a submatrix of $\mathbf{X}$

Because $\mathbf{X}$ has less than full column rank, it is often helpful to use a full column rank sub-matrix of $\mathbf{X}$, having the same rank as $\mathbf{X}$, namely p^*. Call this sub-matrix $\mathbf{X}^*$. It can be any p^* linearly independent columns of $\mathbf{X}$, not necessarily p^* contiguous columns. Despite this, there is no loss of generality if we assume $\mathbf{X}^*$ to be the first p^* columns of $\mathbf{X}$ and partition $\mathbf{X}$ as

$$\mathbf{X} = [\mathbf{X}^* \quad \mathbf{X}^o] \quad \text{for } \mathbf{X}^*_{N\times p^*} \text{ of rank } p^* \,. \tag{2.58}$$

Now define $\mathbf{M}^*$, $\mathbf{P}^*$ and $\mathbf{S}^*$ as $\mathbf{M}$, $\mathbf{P}$ and $\mathbf{S}$ with $\mathbf{X}$ replaced by $\mathbf{X}^*$. Then

$$\mathbf{M} = \mathbf{M}^*, \quad \mathbf{P} = \mathbf{P}^* \text{ and } \mathbf{S} = \mathbf{S}^* \,. \tag{2.59}$$

We prove (2.59) for $\mathbf{P} = \mathbf{P}^*$. For $\mathbf{L}$ of $\mathbf{V}^{-1} = \mathbf{L}'\mathbf{L}$ in (1.16), the partitioning (2.58) gives

$$\mathbf{LX} = [\mathbf{LX}^* \quad \mathbf{LX}^o] \equiv [\mathbf{Y}^* \quad \mathbf{Y}^o] \equiv \mathbf{Y} \,,$$

so defining $\underset{\sim}{Y}^*$ and $\underset{\sim}{Y}^o$. Then $\underset{\sim}{P}$ of (2.26) is

$$\begin{aligned}
\underset{\sim}{P} &= \underset{\sim}{L}'\left\{\underset{\sim}{I} - [\underset{\sim}{Y}^* \quad \underset{\sim}{Y}^o]\begin{bmatrix} \underset{\sim}{Y}^{*\prime}\underset{\sim}{Y}^* & \underset{\sim}{Y}^{*\prime}\underset{\sim}{Y}^o \\ \underset{\sim}{Y}^{o\prime}\underset{\sim}{Y}^* & \underset{\sim}{Y}^{o\prime}\underset{\sim}{Y}^o \end{bmatrix}^- \begin{bmatrix} \underset{\sim}{Y}^{*\prime} \\ \underset{\sim}{Y}^{o\prime} \end{bmatrix}\right\}\underset{\sim}{L} \\
&= \underset{\sim}{L}'\left\{\underset{\sim}{I} - [\underset{\sim}{Y}^* \quad \underset{\sim}{Y}^o]\begin{bmatrix} (\underset{\sim}{Y}^{*\prime}\underset{\sim}{Y}^*)^{-1} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} \begin{bmatrix} \underset{\sim}{Y}^{*\prime} \\ \underset{\sim}{Y}^{o\prime} \end{bmatrix}\right\}\underset{\sim}{L} \\
&= \underset{\sim}{L}'[\underset{\sim}{I} - \underset{\sim}{Y}^*(\underset{\sim}{Y}^{*\prime}\underset{\sim}{Y}^*)^{-1}\underset{\sim}{Y}^{*\prime}]\underset{\sim}{L} \\
&= \underset{\sim}{L}'\underset{\sim}{L} - \underset{\sim}{L}'\underset{\sim}{L}\underset{\sim}{X}^*(\underset{\sim}{X}^{*\prime}\underset{\sim}{L}'\underset{\sim}{L}\underset{\sim}{X}^*)^{-1}\underset{\sim}{X}^{*\prime}\underset{\sim}{L}'\underset{\sim}{L} \\
&= \underset{\sim}{V}^{-1} - \underset{\sim}{V}^{-1}\underset{\sim}{X}^*(\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^*)^{-1}\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1} \\
&= \underset{\sim}{P}^* \, .
\end{aligned} \tag{2.60}$$

The other two results in (2.50) are established in similar fashion. Furthermore, (2.50) holds true for $\underset{\sim}{X}^*$ being <u>any</u> p^* linearly independent columns of $\underset{\sim}{X}$, not just the first p^* such columns as the partitioning (2.49) would indicate.

d. The Partitioned Matrices $\underset{\sim}{B}$ and $\underset{\sim}{C}$

The matrices

$$\underset{\sim}{B} \equiv \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{Z} \\ \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z} \end{bmatrix} = \underset{\sim}{B}' \tag{2.61}$$

and

$$\underset{\sim}{C} \equiv \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D} \\ \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D} \end{bmatrix} = \underset{\sim}{B}\begin{bmatrix} \underset{\sim}{I} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{D} \end{bmatrix} \tag{2.62}$$

play important roles in what are called the mixed model equations (MME's). As a result, we are interested in generalized inverses of B and C.

Note first, that B is symmetric and requires D to be non-singular. In contrast, C is not symmetric and exists whether D is singular or not.

The existence of certain forms of generalized inverses of partitioned matrices is considered at length by Marsaglia and Styan [1974a, b]. Their results show that the existence of $(I + Z'R^{-1}ZD)^{-1} = T^{*}$ in (2.16), whether D is singular or non-singular, ensures that the following expression is a generalized inverse of C:

$$C^{-} = \begin{bmatrix} 0 & 0 \\ 0 & T^{*} \end{bmatrix} + \begin{bmatrix} I \\ -T^{*}Z'R^{-1}X \end{bmatrix} Y^{-}[I \quad -X'R^{-1}ZDT^{*}]$$

for

$$\begin{aligned} Y &= X'R^{-1}X - X'R^{-1}ZD(I + Z'R^{-1}ZD)^{-1}Z'R^{-1}X \\ &= X'[R^{-1} - R^{-1}ZD(I + Z'R^{-1}ZD)^{-1}Z'R^{-1}]X \\ &= X'V^{-1}X \text{ , from (2.14).} \end{aligned}$$

Hence

$$C^{-} = \begin{bmatrix} 0 & 0 \\ 0 & T^{*} \end{bmatrix} + \begin{bmatrix} I \\ -T^{*}Z'R^{-1}X \end{bmatrix} (X'V^{-1}X)^{-}[I \quad -X'R^{-1}ZDT^{*}] \text{ .} \tag{2.63}$$

Another generalized inverse of C is

$$C^{\sim} = \begin{bmatrix} (X'R^{-1}X)^{-} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -(X'R^{-1}X)^{-}X'R^{-1}ZD \\ I \end{bmatrix} W^{-}[-Z'R^{-1}X(X'R^{-1}X)^{-} \quad I]$$

for

$$\begin{aligned}\mathbf{W} &= \mathbf{I} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D} - \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D} \\ &= \mathbf{I} + \mathbf{Z}'[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{R}^{-1}]\mathbf{Z}\mathbf{D} \\ &= \mathbf{I} + \mathbf{Z}'\mathbf{S}\mathbf{Z}\mathbf{D}\ , \\ &= \mathbf{T}^{-1} \text{ of } (2.47).\end{aligned}$$

Hence

$$\mathbf{C}^{\sim} = \begin{bmatrix} (\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D} \\ \mathbf{I} \end{bmatrix} \mathbf{T}[-\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-} \quad \mathbf{I}]\ . \qquad (2.64)$$

Note that just as $\mathbf{C}$ exists whether $\mathbf{D}$ is singular or not, so also do $\mathbf{C}^{-}$ and $\mathbf{C}^{\sim}$.

In contrast, $\mathbf{B}$ of (2.61) does not exist if $\mathbf{D}^{-1}$ does not, a point which is further emphasized by writing, from (2.62),

$$\mathbf{B} = \mathbf{C}\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{bmatrix}$$

when $\mathbf{D}^{-1}$ exists. Generalized inverses of $\mathbf{B}$ corresponding to (2.63) and (2.64) are then

$$\mathbf{B}^{-} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}\mathbf{C}^{-} \quad \text{and} \quad \mathbf{B}^{\sim} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}\mathbf{C}^{\sim}\ .$$

With $\mathbf{D}^{-1}$ existing these are

$$\mathbf{B}^{-} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}\mathbf{T}^{*} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -\mathbf{D}\mathbf{T}^{*}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} \end{bmatrix} (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}[\mathbf{I} \quad -\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D}\mathbf{T}^{*}] \qquad (2.65)$$

where

$$\mathbf{DT}^* = \mathbf{D}(\mathbf{I} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D})^{-1} = (\mathbf{D}^{-1} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z})^{-1} \; ; \tag{2.66}$$

and

$$\mathbf{B}^{\sim} = \begin{bmatrix} (\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^- & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D} \\ \mathbf{D} \end{bmatrix} \mathbf{T}[-\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^- \quad \mathbf{I}]$$

$$= \begin{bmatrix} (\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^- & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{I} \end{bmatrix} \mathbf{DT}[-\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^- \quad \mathbf{I}] \tag{2.67}$$

where

$$\mathbf{DT} = \mathbf{D}(\mathbf{I} + \mathbf{Z}'\mathbf{SZD})^{-1} = (\mathbf{D}^{-1} + \mathbf{Z}'\mathbf{SZ})^{-1} \; . \tag{2.68}$$

We make extensive use of these results in Chapter 3.

e. Matrices $\mathbf{K}'$ satisfying $\mathbf{K}'\mathbf{X} = \mathbf{0}$

Matrices $\mathbf{K}'$ for which $\mathbf{K}'\mathbf{X} = \mathbf{0}$ play an important role in REML, because it is based on using linear functions of the observations, $\mathbf{K}'\mathbf{y}$, that do not involve the fixed effects: $\mathbf{K}'\mathbf{y} = \mathbf{K}'\mathbf{X}\boldsymbol{\alpha} + \mathbf{K}'(\mathbf{Zb} + \mathbf{e}) = \mathbf{K}'(\mathbf{Zb} + \mathbf{e})$. Properties of such matrices $\mathbf{K}'$ are now given.

Lemma 2.1. For $\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ of (2.17), a matrix $\mathbf{K}'$ satisfies $\mathbf{K}'\mathbf{X} = \mathbf{0}$ if and only if

$$\mathbf{K}' = \mathbf{W}'\mathbf{M} \quad \text{for some } \mathbf{W}' \; . \tag{2.69}$$

Proof: If $\mathbf{K}' = \mathbf{W}'\mathbf{M}$, then $\mathbf{K}'\mathbf{X} = \mathbf{W}'\mathbf{MX} = \mathbf{0}$, by (2.19).

If $\mathbf{K}'\mathbf{X} = \mathbf{0}$, then $\mathbf{X}'\mathbf{k} = \mathbf{0}$ for $\mathbf{k}$ being any column of $\mathbf{K}$. Therefore from linear equation theory (e.g., Searle [1971b, p. 13, Corollary])

$$\mathbf{k} = (\mathbf{I} - \mathbf{X}'^-\mathbf{X}')\mathbf{w} \tag{2.70}$$

for some arbitrary vector $\underset{\sim}{w}$. From (2.9), a generalized inverse $\underset{\sim}{X}'^{-}$ of $\underset{\sim}{X}'$ is $\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}$ so that (2.70) is

$$\underset{\sim}{k} = [\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}']\underset{\sim}{w} = \underset{\sim}{M}\underset{\sim}{w} . \tag{2.71}$$

Hence $\underset{\sim}{K} = \underset{\sim}{W}'\underset{\sim}{M}$. <u>Q.E.D.</u>

Linear equation theory also indicates that there are no more than $N - p^*$ linearly independent (LIN) solutions $\underset{\sim}{k}$ obtainable from (2.70). Hence the maximum number of LIN rows in $\underset{\sim}{K}'$ is $N - p^*$. Since the use of $\underset{\sim}{K}'$ is to be through $\underset{\sim}{K}'\underset{\sim}{y}$, there will be no merit in having some elements of $\underset{\sim}{K}'\underset{\sim}{y}$ being linear combinations of others. We therefore confine attention to $\underset{\sim}{K}'$ of full row rank, embracing a set of $N - p^*$ LIN solutions to (2.70). This means we restrict ourselves to matrices $\underset{\sim}{K}'$ of the form

$$(\underset{\sim}{K}')_{N-p^*\times N}, \text{ of full row rank } N - p^*, \text{ such that } \underset{\sim}{K}'\underset{\sim}{X} = \underset{\sim}{0} . \tag{2.72}$$

This also means that

$$\underset{\sim}{K}' = \underset{\sim}{W}'\underset{\sim}{M} \quad \text{for } \underset{\sim}{W}' \text{ of full row rank } N - p^* . \tag{2.73}$$

Lemma 2.2. $\underset{\sim}{x}'\underset{\sim}{y} = 0$, for real non-null vectors $\underset{\sim}{x}$ and $\underset{\sim}{y}$, implies that $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are LIN.

Proof: Assume $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are not LIN so that for non-zero λ_1 and λ_2, $\lambda_1\underset{\sim}{x} + \lambda_2\underset{\sim}{y} = 0$. Then $\lambda_1\underset{\sim}{x}'\underset{\sim}{x} + \lambda_2\underset{\sim}{x}'\underset{\sim}{y} = 0$, i.e., $\lambda_1\underset{\sim}{x}'\underset{\sim}{x} = 0$, giving $\lambda_1 = 0$, which contradicts the assumption. <u>Q.E.D.</u>

Lemma 2.3. $[\underset{\sim}{K} \quad \underset{\sim}{X}^*]$ is non-singular.

Proof: $\underset{\sim}{K}'\underset{\sim}{X} = \underset{\sim}{0}$ implies, by Lemma 1, that rows of $\underset{\sim}{K}'$ and columns of $\underset{\sim}{X}^*$ are LIN. Hence in $[\underset{\sim}{K} \quad \underset{\sim}{X}^*]$ the $N - p^* + p^* = N$ columns, of order N, are LIN, and so $[\underset{\sim}{K} \quad \underset{\sim}{X}^*]$ is non-singular. <u>Q.E.D.</u>

Lemma 2.4. $\underset{\sim}{K}' = \underset{\sim}{F}'\underset{\sim}{A}'$ for some non-singular $\underset{\sim}{F}'$. (2.74)

Proof: From (2.73), $\underset{\sim}{K} = \underset{\sim}{W}'\underset{\sim}{M}$, giving $\underset{\sim}{K}' = \underset{\sim}{W}'\underset{\sim}{A}\underset{\sim}{A}'$ from (2.22). Hence $\underset{\sim}{K}' = \underset{\sim}{F}'\underset{\sim}{A}'$ for $\underset{\sim}{F}' = \underset{\sim}{W}'\underset{\sim}{A}$, square of order $N - p^*$. Therefore $N - p^* = r(\underset{\sim}{K}') \leq r(\underset{\sim}{F}')$ and so $r(\underset{\sim}{F}') = N - p^*$ and hence (2.74) is true. Q.E.D.

Lemma 2.5. $\underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K}' = \underset{\sim}{A}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{A} = \underset{\sim}{P}$. (2.75)

Proof: The second equality comes from (2.35). The first is true because, from (2.74), $\underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K}' = \underset{\sim}{A}\underset{\sim}{F}(\underset{\sim}{F}'\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{F})^{-1}\underset{\sim}{F}'\underset{\sim}{A}' = \underset{\sim}{A}\underset{\sim}{F}\underset{\sim}{F}^{-1}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{F}'^{-1}\underset{\sim}{F}'\underset{\sim}{A}' = \underset{\sim}{A}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{A}'$. Q.E.D.

Lemma 2.6. $\underset{\sim}{A}'$ is a possible $\underset{\sim}{K}'$.

Proof: Compare (2.72) and (2.24). Q.E.D.

Corollary: $[\underset{\sim}{A} \quad \underset{\sim}{X}^*]$ is non-singular.

Proof: See Lemma 2.3. Q.E.D.

Lemma 2.7.

$$|\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{A}'| = \frac{|\underset{\sim}{V}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^*|}{|\underset{\sim}{X}^{*\prime}\underset{\sim}{X}^*|} . \tag{2.76}$$

Proof:

$$\begin{bmatrix} \underset{\sim}{A}' \\ \underset{\sim}{X}^{*\prime} \end{bmatrix} \underset{\sim}{V} [\underset{\sim}{A} \quad \underset{\sim}{X}^*] = \begin{bmatrix} \underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A} & \underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{X}^* \\ \underset{\sim}{X}^{*\prime}\underset{\sim}{V}\underset{\sim}{A} & \underset{\sim}{X}^{*\prime}\underset{\sim}{V}\underset{\sim}{X}^* \end{bmatrix} .$$

Taking determinants of both sides and using (2.5) gives

$$|\underset{\sim}{V}| \begin{vmatrix} \underset{\sim}{A}'\underset{\sim}{A} & \underset{\sim}{A}'\underset{\sim}{X}^* \\ \underset{\sim}{X}^{*\prime}\underset{\sim}{A} & \underset{\sim}{X}^{*\prime}\underset{\sim}{X}^* \end{vmatrix} = |\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{V}\underset{\sim}{X}^* - \underset{\sim}{X}^{*\prime}\underset{\sim}{V}\underset{\sim}{A}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{X}^*|$$

and on using $\underset{\sim}{A}'\underset{\sim}{A} = \underset{\sim}{I}$ from (2.22), $\underset{\sim}{A}'\underset{\sim}{X} = \underset{\sim}{0}$ from (2.24), and hence $\underset{\sim}{A}'\underset{\sim}{X}^* = \underset{\sim}{0}$, this is

$$|\underset{\sim}{V}| \begin{vmatrix} \underset{\sim}{I} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{X}^{*\prime}\underset{\sim}{X}^* \end{vmatrix} = |\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{V}\{\underset{\sim}{V}^{-1} - \underset{\sim}{A}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{A}'\}\underset{\sim}{V}\underset{\sim}{X}^*| \ ;$$

i.e.,

$$|\underset{\sim}{V}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{X}^{*}| = |\underset{\sim}{A}'\underset{\sim}{VA}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{V}(\underset{\sim}{V}^{-1} - \underset{\sim}{P})\underset{\sim}{VX}^{*}|\ , \text{ from } (2.35)$$

$$= |\underset{\sim}{A}'\underset{\sim}{VA}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{V}(\underset{\sim}{V}^{-1} - \underset{\sim}{P}^{*})\underset{\sim}{VX}^{*}|, \text{ from } (2.60)$$

$$= |\underset{\sim}{A}'\underset{\sim}{VA}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{VV}^{-1}\underset{\sim}{X}^{*}(\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^{*})^{-1}\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{VX}^{*}|$$

$$= |\underset{\sim}{A}'\underset{\sim}{VA}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{X}^{*}|^{2} .$$

(2.76) follows at once. Q.E.D.

Lemma 2.8. $|\underset{\sim}{K}'\underset{\sim}{VK}| = |\underset{\sim}{F}|^2|\underset{\sim}{A}'\underset{\sim}{VA}|$.

Proof: $|\underset{\sim}{K}'\underset{\sim}{VK}| = |\underset{\sim}{F}'\underset{\sim}{A}'\underset{\sim}{VAF}| = |\underset{\sim}{F}|^2|\underset{\sim}{A}'\underset{\sim}{VA}|$. Q.E.D.

Lemma 2.9. For any matrix $\underset{\sim}{Q}_{N\times q}$ not of full column rank but with

$$r(\underset{\sim}{X}) + r(\underset{\sim}{Q}) = N \quad \text{and} \quad \underset{\sim}{Q}'\underset{\sim}{X} = \underset{\sim}{0}$$

$$\underset{\sim}{Q}(\underset{\sim}{Q}'\underset{\sim}{VQ})^{-}\underset{\sim}{Q}' = \underset{\sim}{P} .$$

This is a generalization of Lemma 5 and is also a generalization (involving a generalized inverse of $\underset{\sim}{Q}'\underset{\sim}{VQ}$ rather than a regular inverse) of a theorem in Khatri [1966].

Proof: $r(\underset{\sim}{Q}) = N - r(\underset{\sim}{X}) = N - p^{*}$. Because $\underset{\sim}{Q}'\underset{\sim}{X} = \underset{\sim}{0}$ and $\underset{\sim}{Q}'$ has more than $N - p^{*}$ rows (since $\underset{\sim}{Q}$ does not have full column rank), partition $\underset{\sim}{Q}'$ as

$$\underset{\sim}{Q}' = \begin{bmatrix} \underset{\sim}{K}' \\ \underset{\sim}{TK}' \end{bmatrix}$$

for some matrix $\underset{\sim}{T}$, and where $\underset{\sim}{K}'$ has the form described in (2.72). Then

$$\underset{\sim}{Q}(\underset{\sim}{Q}'\underset{\sim}{V}\underset{\sim}{Q})^{-}\underset{\sim}{Q}' = [\underset{\sim}{K} \quad \underset{\sim}{K}\underset{\sim}{T}'] \begin{bmatrix} \underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K} & \underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K}\underset{\sim}{T}' \\ \underset{\sim}{T}\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K} & \underset{\sim}{T}\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K}\underset{\sim}{T}' \end{bmatrix}^{-} \begin{bmatrix} \underset{\sim}{K}' \\ \underset{\sim}{T}\underset{\sim}{K}' \end{bmatrix}$$

$$= [\underset{\sim}{K} \quad \underset{\sim}{K}\underset{\sim}{T}'] \begin{bmatrix} (\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} \begin{bmatrix} \underset{\sim}{K}' \\ \underset{\sim}{T}\underset{\sim}{K}' \end{bmatrix}$$

$$= \underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K}'$$

$$= \underset{\sim}{P} \text{ , from (2.75).} \qquad \underline{\text{Q.E.D.}}$$

2.7. DIFFERENTIATING GENERALIZED INVERSES

a. Differentiable and non-differentiable generalized inverses

Some generalized inverses of a matrix are differentiable and some are not. Consider $\underset{\sim}{W}$ having elements that are differentiable functions of the scalar argument t. If $\underset{\sim}{W}_{11}$ is the leading r × r sub-matrix of $\underset{\sim}{W}$ and is non-singular, with r being the rank of $\underset{\sim}{W}$, then $\underset{\sim}{W}_{11}^{-1}$ is differentiable with respect to t and so is the generalized inverse

$$\underset{\sim}{G} = \begin{bmatrix} \underset{\sim}{W}_{11}^{-1} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} .$$

In contrast, if for $\underset{\sim}{P}$ and $\underset{\sim}{Q}$ being products of elementary operators the equivalent canonical form of $\underset{\sim}{W}$ is

$$\underset{\sim}{P}\underset{\sim}{W}\underset{\sim}{Q} = \begin{bmatrix} \underset{\sim}{D}_r & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix}, \quad \text{then} \quad \underset{\sim}{G}_2 = \underset{\sim}{Q} \begin{bmatrix} \underset{\sim}{D}_r^{-1} & \underset{\sim}{X} \\ \underset{\sim}{Y} & \underset{\sim}{Z} \end{bmatrix} \underset{\sim}{P} \tag{2.77}$$

is a generalized inverse of $\underset{\sim}{W}$ for any matrices $\underset{\sim}{X}$, $\underset{\sim}{Y}$ and $\underset{\sim}{Z}$; and if the latter are not differentiable, then neither is $\underset{\sim}{G}_2$.

Providing $\underset{\sim}{G}$ is differentiable, differentiating its defining equation $\underset{\sim}{W}\underset{\sim}{G}\underset{\sim}{W} = \underset{\sim}{W}$ gives

$$(\partial \underset{\sim}{W})\underset{\sim}{G}\underset{\sim}{W} + \underset{\sim}{W}(\partial \underset{\sim}{G})\underset{\sim}{W} + \underset{\sim}{W}\underset{\sim}{G}(\partial \underset{\sim}{W}) = \partial \underset{\sim}{W} . \qquad (2.78)$$

The provision that $\underset{\sim}{G}$ be differentiable for (2.78) to exist, was brought to our attention by Harville (personal communication). It is required because, as is illustrated by $\underset{\sim}{G}_2$ of (2.77), $\underset{\sim}{W}\underset{\sim}{G}\underset{\sim}{W} = \underset{\sim}{W}$ does not of itself ensure differentiability of $\underset{\sim}{G}$.

When $\underset{\sim}{G}$ is differentiable, (2.78) exists, whereupon, pre- and post-multiplying it by $\underset{\sim}{W}\underset{\sim}{G}$ and $\underset{\sim}{G}\underset{\sim}{W}$ respectively leads to

$$\underset{\sim}{W}(\partial \underset{\sim}{G})\underset{\sim}{W} = -\underset{\sim}{W}\underset{\sim}{G}(\partial \underset{\sim}{W})\underset{\sim}{G}\underset{\sim}{W} . \qquad (2.79)$$

This is the generalized inverse analogue of the regular inverse result $\partial \underset{\sim}{W}^{-1} = -\underset{\sim}{W}^{-1}(\partial \underset{\sim}{W})\underset{\sim}{W}^{-1}$, to which it and (2.78) reduce when $\underset{\sim}{W}$ is non-singular.

b. Differentiating $\underset{\sim}{P}$

We will have occasion to want to differentiate $\underset{\sim}{P}$. Using its definition (2.26), which involves $(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^-$, we would have to be concerned about the differentiability of whatever generalized inverse were to be used for $(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^-$. But this issue can be avoided by using (2.35):

$$\underset{\sim}{P} = \underset{\sim}{A}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{A} .$$

Then

$$\partial \underset{\sim}{P} = -\underset{\sim}{A}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{A}'\partial \underset{\sim}{V}\underset{\sim}{A}(\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A})^{-1}\underset{\sim}{A} = -\underset{\sim}{P}(\partial \underset{\sim}{V})\underset{\sim}{P} . \qquad (2.80)$$

2.8. DIFFERENTIATING $\underset{\sim}{V}$

Because, as in (1.18), $\underset{\sim}{V} = \sum_{i=1}^{c} \underset{\sim}{Z}_i\underset{\sim}{Z}_i'\sigma_i^2 + \sigma_0^2\underset{\sim}{I}$ it is clear that

$$\frac{\partial \underset{\sim}{V}}{\partial \sigma_0^2} = \underset{\sim}{I} \quad \text{and} \quad \frac{\partial \underset{\sim}{V}}{\partial \sigma_i^2} = \underset{\sim}{Z}_i \underset{\sim}{Z}_i', \qquad i = 1, \cdots, c \tag{2.81}$$

as indicated in (2.3). And for the change of variables $\gamma_0 = \sigma_0^2$ and $\gamma_i = \sigma_i^2/\sigma_0^2$, equivalent to

$$\sigma_0^2 = \gamma_0 \quad \text{and} \quad \sigma_i^2 = \gamma_0 \gamma_i \quad \text{for } i = 1, \cdots, c, \tag{2.82}$$

$$\underset{\sim}{V} = \gamma_0 \underset{\sim}{H} \quad \text{for} \quad \underset{\sim}{H} = \sum_{i=1}^{c} \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \gamma_i + \underset{\sim}{I} \tag{2.83}$$

as in (1.21) and (1.22). Therefore

$$\frac{\partial \underset{\sim}{V}}{\partial \gamma_0} = \underset{\sim}{H} \quad \text{and} \quad \frac{\partial \underset{\sim}{V}}{\partial \gamma_i} = \gamma_0 \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \text{ for } i = 1, \cdots, c. \tag{2.84}$$

Since $\gamma_0 \equiv \sigma_0^2$, the consistency of the first equations in each of (2.81) and (2.84) warrants demonstration. From elementary calculus we know that

$$\frac{\partial \underset{\sim}{V}}{\partial \gamma_i} = \sum_{k=0}^{c} \frac{\partial \underset{\sim}{V}}{\partial \sigma_k^2} \frac{\partial \sigma_k^2}{\partial \gamma_i}$$

and for i = 0 this gives, using (2.81) and (2.82),

$$\frac{\partial \underset{\sim}{V}}{\partial \gamma_0} = \underset{\sim}{I} + \sum_{k=1}^{c} \underset{\sim}{Z}_k \underset{\sim}{Z}_k' \gamma_k = \underset{\sim}{H}, \tag{2.85}$$

from (2.83); and so the first result in (2.84) is established. Similarly,

$$\frac{\partial \underset{\sim}{V}}{\partial \gamma_i} = \underset{\sim}{0} + \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \gamma_0 = \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \gamma_0, \qquad \text{for } i = 1, \cdots, c, \tag{2.86}$$

which is the second result in (2.84). It is because (2.82) represents not just the substitution $\gamma_0 = \sigma_0^2$ but rather the complete change of variables from $\dot{\underset{\sim}{\sigma}}$ to $\dot{\underset{\sim}{\gamma}}$ [in terms of (1.27)], that the apparent inconsistency of $\partial \underset{\sim}{V}/\partial\sigma_0^2 = \underset{\sim}{I}$ and $\partial\underset{\sim}{V}/\partial\gamma_0 = \underset{\sim}{H}$ occurs.

2.9. THE VEC AND VECH OPERATORS

Operators that put elements of a matrix into a single vector are useful in a theorem (Sec. 2.10) that is the basis of MINQUE.

a. Definitions

The vec of any matrix $\underset{\sim}{X}$, to be denoted by vec$\underset{\sim}{X}$, is the vector formed by stacking the columns of $\underset{\sim}{X}$ one under another in a single column. Thus for $\underset{\sim}{X}$ of order $r \times c$ the order of vec$\underset{\sim}{X}$ is $rc \times 1$. For example, with

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 2 & 7 \\ 3 & 9 & 8 \\ 6 & 4 & 5 \end{bmatrix}, \quad \text{vec}\underset{\sim}{X} = [1\ 3\ 6\ 2\ 9\ 4\ 7\ 8\ 5]' .$$

Note that $(\text{vec}\underset{\sim}{X})'$ and vec$\underset{\sim}{X}'$ are different: vec$\underset{\sim}{X}'$ is the vec of the matrix $\underset{\sim}{X}'$ which, in the example, is

$$\text{vec}\underset{\sim}{X}' = [1\ 2\ 7\ 3\ 9\ 8\ 6\ 4\ 5]' .$$

The vector denoted by vech$\underset{\sim}{X}$ is similar to vec$\underset{\sim}{X}$ except that it is defined only for $\underset{\sim}{X}$ being square and only that part of each column of $\underset{\sim}{X}$ that is on and below the diagonal of $\underset{\sim}{X}$ is put into vech$\underset{\sim}{X}$; for the preceding example,

$$\text{vech}\underset{\sim}{X} = [1\ 3\ 6\ 9\ 4\ 5]' .$$

Neudecker [1969] has recently used vec$\underset{\sim}{X}$ for obtaining some Jacobians, developing useful results concerning products and traces along the way, and Searle [1978] introduces the vech operator for obtaining Jacobians that involve symmetric matrices.

For then, with $\mathbf{X} = \mathbf{X}'$, vech$\mathbf{X}$ represents just the distinctly different elements of $\mathbf{X}$; e.g., for

$$\mathbf{X} = \begin{bmatrix} 7 & 3 & 9 \\ 3 & 2 & 6 \\ 9 & 6 & 8 \end{bmatrix}, \quad \text{vech}\mathbf{X} = [7\ 3\ 9\ 2\ 6\ 8]' .$$

Clearly, for $\mathbf{X}$ of order n, vech$\mathbf{X}$ has order $t_n \times 1$, for

$$t_n = \tfrac{1}{2}n(n+1) .$$

b. <u>Relationships between vec and vech for symmetric matrices</u>

The definitions of vec$\mathbf{X}$ and vech$\mathbf{X}$ mean that for symmetric $\mathbf{X}$ each is a linear transformation of the other. We represent these relationships as

$$\text{vech}\mathbf{X} = \mathbf{H}\text{vec}\mathbf{X} \quad \text{and} \quad \text{vec}\mathbf{X} = \mathbf{G}\text{vech}\mathbf{X} . \tag{2.87}$$

[This $\mathbf{H}$ is not to be confused with $\mathbf{H}$ of $\mathbf{V} = \sigma_Q^2\mathbf{H}$.] Associated with $\mathbf{H}$ and $\mathbf{G}$ of (2.87) is the permuted identity matrix $\mathbf{I}_{(n,n)}$ defined by MacRae [1974]. This is a symmetric matrix of order n^2, partitioned into $n \times n$ square sub-matrices each of order n, with, for i, j = 1, $\cdots$, n, the (i, j)'th sub-matrix having unity for its (j, i)'th element and zeros elsewhere. Examples of $\mathbf{I}_{(n,n)}$, $\mathbf{H}$ and $\mathbf{G}$ for n = 3 are, with dots representing zeros,

$$\mathbf{I}_{(3,3)} = \left[\begin{array}{ccc|ccc|ccc} 1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot \\ \cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot \\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot \\ \hline \cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot \\ \cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot \\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot \\ \hline \cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot \\ \cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot \\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1 \end{array}\right], \quad \mathbf{H}_{t_3 \times 3^2} = \left[\begin{array}{ccc|ccc|ccc} 1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot \\ \cdot&\frac{1}{2}&\cdot&\frac{1}{2}&\cdot&\cdot&\cdot&\cdot&\cdot \\ \cdot&\cdot&\frac{1}{2}&\cdot&\cdot&\cdot&\frac{1}{2}&\cdot&\cdot \\ \hline \cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot \\ \cdot&\cdot&\cdot&\cdot&\cdot&\frac{1}{2}&\cdot&\frac{1}{2}&\cdot \\ \hline \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1 \end{array}\right] \quad \text{and} \quad \mathbf{G}_{3^2 \times t_3} = \left[\begin{array}{ccc|cc|c} 1&\cdot&\cdot&\cdot&\cdot&\cdot \\ \cdot&1&\cdot&\cdot&\cdot&\cdot \\ \cdot&\cdot&1&\cdot&\cdot&\cdot \\ \hline \cdot&1&\cdot&\cdot&\cdot&\cdot \\ \cdot&\cdot&\cdot&1&\cdot&\cdot \\ \cdot&\cdot&\cdot&\cdot&1&\cdot \\ \hline \cdot&\cdot&1&\cdot&\cdot&\cdot \\ \cdot&\cdot&\cdot&\cdot&1&\cdot \\ \cdot&\cdot&\cdot&\cdot&\cdot&1 \end{array}\right] .$$

Although there are numerous relationships between $\underset{\sim}{I}_{n^2}$, $\underset{\sim}{I}_{(n,n)}$, $\underset{\sim}{H}$ and $\underset{\sim}{G}$, we make no particular use of them in this Notebook and so shall not describe them. The interested reader is referred to Roth [1934], Neudecker [1969], Tracy and Dwyer [1969], and MacRae [1974] for details.

c. Trace and vec

Neudecker [1969] develops the following result connecting the trace and vec operators: for any matrices $\underset{\sim}{A}$ and $\underset{\sim}{B}$ for which $\underset{\sim}{A}\underset{\sim}{B}$ exists,

$$\mathrm{tr}(\underset{\sim}{A}\underset{\sim}{B}) = (\mathrm{vec}\underset{\sim}{A}')'\mathrm{vec}\underset{\sim}{B} \ . \tag{2.88}$$

The verity of this is clear: $(\mathrm{vec}\underset{\sim}{A}')'$ is the row vector of elements a_{ij} of $\underset{\sim}{A}$, ordered j within i, and $\mathrm{vec}\underset{\sim}{B}$ is the column vector of elements b_{ji} ordered i within j. Hence

$$(\mathrm{vec}\underset{\sim}{A}')'\mathrm{vec}\underset{\sim}{B} = \sum_{i,j} a_{ij}b_{ji} = \sum_i(\sum_j a_{ij}b_{ji}) = \mathrm{tr}(\underset{\sim}{A}\underset{\sim}{B}) \ .$$

d. The vec of a product

For any matrices $\underset{\sim}{A}$, $\underset{\sim}{B}$ and $\underset{\sim}{C}$ (not just the special matrices defined earlier), Neudecker [1969] has a result for the vec of the product $\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C}$, of suitably conformable matrices

$$\mathrm{vec}(\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C}) = (\underset{\sim}{C}' * \underset{\sim}{A})\mathrm{vec}\underset{\sim}{B} \ . \tag{2.89}$$

Derivation of this result is as follows. The j'th sub-vector of $\mathrm{vec}(\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C})$ is

$$\text{j'th column of } \underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C} = \underset{\sim}{A}\underset{\sim}{B}(\text{j'th column of } \underset{\sim}{C})$$

$$= \sum_i c_{ij}(\text{i'th column of } \underset{\sim}{A}\underset{\sim}{B}) \tag{2.90}$$

$$= \sum_i c_{ij}\underset{\sim}{A}(\text{i'th column of } \underset{\sim}{B}) \tag{2.91}$$

$$= [c_{1j}\underset{\sim}{A} \cdots c_{rj}\underset{\sim}{A}]\begin{bmatrix}\underset{\sim}{b}_1\\ \vdots \\ \underset{\sim}{b}_r\end{bmatrix} \quad \text{for } \underset{\sim}{B} = [\underset{\sim}{b}_1 \cdots \underset{\sim}{b}_r]$$

$$= [\underset{\sim}{c}_j' * \underset{\sim}{A}]\mathrm{vec}\underset{\sim}{B} \quad \text{for } \underset{\sim}{C} = [\underset{\sim}{c}_1 \cdots \underset{\sim}{c}_n] \ .$$

Therefore

$$\mathrm{vec}(\underset{\sim}{ABC}) = \begin{bmatrix} \underset{\sim}{c}_1' * \underset{\sim}{A} \\ \vdots \\ \underset{\sim}{c}_n' * \underset{\sim}{A} \end{bmatrix} \mathrm{vec}\underset{\sim}{B} = (\underset{\sim}{C}' * \underset{\sim}{A})\mathrm{vec}\underset{\sim}{B} .$$

Three special cases of (2.89) are, for any conformable $\underset{\sim}{P}$ and $\underset{\sim}{Q}$,

$$\mathrm{vec}(\underset{\sim}{PQ}) = (\underset{\sim}{I} * \underset{\sim}{P})\mathrm{vec}\underset{\sim}{Q} = (\underset{\sim}{Q}' * \underset{\sim}{I})\mathrm{vec}\underset{\sim}{P} = (\underset{\sim}{Q}' * \underset{\sim}{P})\mathrm{vec}\underset{\sim}{I} , \tag{2.92}$$

obtained by using $\underset{\sim}{P}, \underset{\sim}{Q}$ respectively for the symbols $\underset{\sim}{A}, \underset{\sim}{B}$ and $\underset{\sim}{B}, \underset{\sim}{C}$ and $\underset{\sim}{A}, \underset{\sim}{C}$ in turn in (2.89). Another special case is

$$\mathrm{vec}(\underset{\sim}{yy}') = \underset{\sim}{y} * \underset{\sim}{y} . \tag{2.93}$$

2.10. A MINIMIZATION THEOREM

Theorem: The symmetric matrix $\underset{\sim}{Q}$ with $\underset{\sim}{QX} = \underset{\sim}{0}$ which minimizes $\mathrm{tr}(\underset{\sim}{Q}^2)$ subject to

$$\mathrm{tr}(\underset{\sim}{QW}_i) = t_i \quad \text{for } i = 1, \cdots, p \text{ with } \underset{\sim}{W}_i \text{ symmetric} \tag{2.94}$$

is

$$\underset{\sim}{Q} = \sum_{i=1}^{p} \lambda_i \underset{\sim}{MW}_i\underset{\sim}{M} \tag{2.95}$$

where

$$\sum_{i=1}^{p} \lambda_i \mathrm{tr}(\underset{\sim}{MW}_i\underset{\sim}{MW}_j) = t_i \;;\; \text{i.e., } \{\mathrm{tr}(\underset{\sim}{MW}_i\underset{\sim}{MW}_j)\}\underset{\sim}{\lambda} = \underset{\sim}{t} \tag{2.96}$$

for $i, j = 1, \cdots, p$, and for $\underset{\sim}{M} = \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'$ of (2.17).

This theorem is due to Rao [1971a]. We give a proof that uses the vec and vech operators of the preceding section.

Proof: Because $\underset{\sim}{QX} = \underset{\sim}{0}$, it is known from (2.69) that $\underset{\sim}{Q} = \underset{\sim}{EM}$ for some matrix $\underset{\sim}{E}$;

and because $\underset{\sim}{Q}$ is symmetric, $\underset{\sim}{Q} = \underset{\sim}{Q}' = \underset{\sim}{E}\underset{\sim}{M} = \underset{\sim}{M}'\underset{\sim}{E}' = \underset{\sim}{M}\underset{\sim}{E}'$. Therefore, on multiplying $\underset{\sim}{E}\underset{\sim}{M} = \underset{\sim}{M}\underset{\sim}{E}'$ by $\underset{\sim}{M}$ we get $\underset{\sim}{E}\underset{\sim}{M}^2 = \underset{\sim}{M}\underset{\sim}{E}'\underset{\sim}{M}$, i.e., $\underset{\sim}{E}\underset{\sim}{M} = \underset{\sim}{M}\underset{\sim}{E}'\underset{\sim}{M}$. But $\underset{\sim}{E}\underset{\sim}{M} = \underset{\sim}{Q}$ and $\underset{\sim}{Q}$ is symmetric, so that

$$\underset{\sim}{Q} = \underset{\sim}{M}\underset{\sim}{E}\underset{\sim}{M} \quad \text{for some symmetric matrix } \underset{\sim}{E}\ , \tag{2.97}$$

a result noted by LaMotte [1973].

To minimize $\mathrm{tr}(\underset{\sim}{Q}^2)$ subject to (2.94), let $2\underset{\sim}{\lambda}$ be a vector of Lagrange multipliers and, using (2.97), minimize

$$\Delta = \mathrm{tr}(\underset{\sim}{M}\underset{\sim}{E}\underset{\sim}{M})^2 - 2\sum_{i=1}^{p}\lambda_i[\mathrm{tr}(\underset{\sim}{M}\underset{\sim}{E}\underset{\sim}{M}\underset{\sim}{W}_i) - t_i] \tag{2.98}$$

$$= \mathrm{tr}(\underset{\sim}{E}\underset{\sim}{M}\underset{\sim}{E}\underset{\sim}{M}) - 2\Sigma\lambda_i[\mathrm{tr}(\underset{\sim}{E}\underset{\sim}{M}\underset{\sim}{W}_i\underset{\sim}{M}) - t_i]$$

$$= [\mathrm{vec}(\underset{\sim}{M}'\underset{\sim}{E}')]'\mathrm{vec}(\underset{\sim}{E}\underset{\sim}{M}) - 2\Sigma\lambda_i[(\mathrm{vec}\underset{\sim}{E}')'\mathrm{vec}(\underset{\sim}{M}\underset{\sim}{W}_i\underset{\sim}{M}) - t_i]\ ,$$

on using (2.88). Then, with (2.92) and the symmetry of $\underset{\sim}{E}$ and $\underset{\sim}{M}$, this becomes

$$\Delta = [(\underset{\sim}{I} * \underset{\sim}{M})\mathrm{vec}\underset{\sim}{E}]'(\underset{\sim}{M}' * \underset{\sim}{I})\mathrm{vec}\underset{\sim}{E} - 2\Sigma\lambda_i[(\mathrm{vec}\underset{\sim}{E})'\mathrm{vec}(\underset{\sim}{M}\underset{\sim}{W}_i\underset{\sim}{M}) - t_i]\ ,$$

and, because $\underset{\sim}{E}$ is symmetric we use (2.87) to get

$$\Delta = (\mathrm{vech}\underset{\sim}{E})'\underset{\sim}{G}'(\underset{\sim}{I} * \underset{\sim}{M})(\underset{\sim}{M} * \underset{\sim}{I})\underset{\sim}{G}\mathrm{vech}\underset{\sim}{E} - 2(\mathrm{vech}\underset{\sim}{E})'\underset{\sim}{G}'\Sigma\lambda_i\mathrm{vec}(\underset{\sim}{M}\underset{\sim}{W}_i\underset{\sim}{M}) + 2\Sigma\lambda_i t_i\ ;$$

i.e.,

$$\Delta = (\mathrm{vech}\underset{\sim}{E})'\underset{\sim}{G}'(\underset{\sim}{M} * \underset{\sim}{M})\underset{\sim}{G}\mathrm{vech}\underset{\sim}{E} - 2(\mathrm{vech}\underset{\sim}{E})'\underset{\sim}{G}'\Sigma\lambda_i\mathrm{vec}\underset{\sim}{M}\underset{\sim}{W}_i\underset{\sim}{M} + 2\Sigma\lambda_i t_i\ . \tag{2.99}$$

The object is to choose $\underset{\sim}{E}$ to minimize Δ. We therefore differentiate Δ with respect to elements of $\mathrm{vech}\underset{\sim}{E}$, using the principles of (2.1) to do so:

$$\partial\Delta/\partial(\mathrm{vech}\underset{\sim}{E}) = 2\underset{\sim}{G}'(\underset{\sim}{M} * \underset{\sim}{M})\underset{\sim}{G}\mathrm{vech}\underset{\sim}{E} - 2\underset{\sim}{G}'\Sigma\lambda_i\mathrm{vec}(\underset{\sim}{M}\underset{\sim}{W}_i\underset{\sim}{M})$$

$$= 2\underset{\sim}{G}'[(\underset{\sim}{M} * \underset{\sim}{M})\underset{\sim}{G}\mathrm{vech}\underset{\sim}{E} - \Sigma\lambda_i\mathrm{vec}(\underset{\sim}{M}\underset{\sim}{W}_i\underset{\sim}{M})]\ . \tag{2.100}$$

Before equating to $\mathbf{0}$ to derive $\mathbf{E}$, note that whatever the solution may be,

$$\partial^2\Delta/\partial(\mathrm{vech}\,\mathbf{E})^2 = 2\mathbf{G}'(\mathbf{M} * \mathbf{M})\mathbf{G} \ .$$

This is positive semi-definite, because $\mathbf{G}$ has full column rank, thus ensuring that we will have minimized and not maximized Δ.

A sufficient condition for (2.100) to be $\mathbf{0}$ is

$$(\mathbf{M} * \mathbf{M})\mathbf{G}\,\mathrm{vech}\,\mathbf{E} = \Sigma\lambda_i \mathrm{vec}(\mathbf{M}\mathbf{W}_i\mathbf{M}) \ ,$$

i.e.,

$$(\mathbf{M} * \mathbf{M})\mathrm{vec}\,\mathbf{E} = \mathrm{vec}(\Sigma\lambda_i \mathbf{M}\mathbf{W}_i\mathbf{M}) \ ,$$

or, using (2.89)

$$\mathbf{M}\mathbf{E}\mathbf{M} = \Sigma\lambda_i \mathbf{M}\mathbf{W}_i\mathbf{M} \ ,$$

i.e.,

$$\mathbf{Q} = \Sigma\lambda_i \mathbf{M}\mathbf{W}_i\mathbf{M} \ .$$

This is (2.95) of the theorem. Using it in (2.94) gives

$$\mathrm{tr}(\sum_i \lambda_i \mathbf{M}\mathbf{W}_i\mathbf{M}\mathbf{W}_j) = t_j, \quad \text{for } j = 1 \cdots p \ ,$$

i.e.,

$$\sum_i \mathrm{tr}(\mathbf{M}\mathbf{W}_i\mathbf{M}\mathbf{W}_j)\lambda_i = t_j, \quad \text{for } j = 1 \cdots p$$

which is the first equality of (2.96); in the second, it is restated using the vector form $\boldsymbol{\lambda}$ for the λ_i's. <u>Q.E.D.</u>

Chapter 3

THE MME's: MIXED MODEL EQUATIONS

3.1. THE EQUATIONS

a. Definition

In the model $\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\mathbf{b} + \mathbf{e}$ of (1.2) with $\text{var}(\mathbf{y}) = \mathbf{V} = \mathbf{ZDZ}' + \mathbf{R}$ of (1.12), a generalized least square (Aitken) solution for $\boldsymbol{\alpha}$ is a solution to the normal equations

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\hat{\boldsymbol{\alpha}} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \, . \tag{3.1}$$

If, instead of representing random effects, $\mathbf{b}$ were to represent fixed effects, the normal equations would be

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}^{o} \\ \mathbf{b}^{o} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} . \tag{3.2}$$

Equations which have come to be called the mixed model equations (MME), and which were first considered by Henderson in Henderson et al. [1959], are (3.2) adapted by adding $\mathbf{D}^{-1}$ to $\mathbf{Z}'\mathbf{RZ}$:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{D}^{-1} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\alpha}} \\ \tilde{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} . \tag{3.3}$$

An equivalent form of these considered by Harville [1977] is

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{ZD} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{I} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{ZD} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\alpha}} \\ \tilde{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \tag{3.4}$$

and

$$\underset{\sim}{D}\underset{\sim}{\tilde{\nu}} = \underset{\sim}{\tilde{b}} \,. \tag{3.5}$$

b. Consistency

Both sets of equations are consistent. In (3.3), which requires that $\underset{\sim}{D}$ be non-singular,

$$(\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}) = \underset{\sim}{D}^{-\frac{1}{2}}(\underset{\sim}{I} + \underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}^{\frac{1}{2}})\underset{\sim}{D}^{-\frac{1}{2}}$$

which is readily established as non-singular using arguments similar to those used for deriving (2.16). Hence (3.3) can be reduced to

$$\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X}\underset{\sim}{\tilde{\alpha}} + \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}(\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}}) = \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{y} \,,$$

which is clearly consistent. Similarly in (3.4), $\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}$ is non-singular, as in (2.49), and the equations can be reduced to

$$\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X}\underset{\sim}{\tilde{\alpha}} + \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}}) = \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{y} \,,$$

which is also consistent.

c. Singular $\underset{\sim}{D}$

Equations (3.3) require $\underset{\sim}{D}$ to be non-singular whereas (3.4) do not. By nature of its definition (1.11), $\underset{\sim}{D}$ is customarily non-singular because, although σ_i^2 is defined for $\sigma_i^2 \geq 0$, models are usually defined only with $\sigma_i^2 > 0$. But if equations (3.3) are used in any iterative computing procedure where computed $\underset{\sim}{D}$ is in terms of computed estimates of the σ_i^2's, any estimated σ_i^2 computed as negative or zero and accordingly given the value 0, as is often the practice, will make the computed $\underset{\sim}{D}$ singular. This would make (3.3) unusable in a computing context, whereas it would not affect (3.4). This is the advantage of (3.4). We therefore consider both sets of equations.

3.2. SOME MATRIX EQUALITIES

Before considering solutions to (3.3) and (3.4) we collect and develop a number of equalities that are useful in deriving the solutions and in subsequently using them.

First, from (2.16) and (2.66)

$$\underset{\sim}{T}^* = (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D})^{-1}, \quad \underset{\sim}{T}^{*-1} - \underset{\sim}{I} = \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D} \quad \text{and} \quad \underset{\sim}{D}\underset{\sim}{T}^* = (\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z})^{-1}, \tag{3.6}$$

the latter when $\underset{\sim}{D}^{-1}$ exists. Also, $\underset{\sim}{D}\underset{\sim}{T}^{*-1}$ is symmetric so that $\underset{\sim}{D}\underset{\sim}{T}^{*-1} = \underset{\sim}{T}^{*-1'}\underset{\sim}{D}$, so giving

$$\underset{\sim}{D}\underset{\sim}{T}^* = \underset{\sim}{T}^{*'}\underset{\sim}{D} . \tag{3.7}$$

Similarly, from (2.47) and (2.68)

$$\underset{\sim}{T} = (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{S}\underset{\sim}{Z}\underset{\sim}{D})^{-1}, \quad \underset{\sim}{T}^{-1} - \underset{\sim}{I} = \underset{\sim}{Z}'\underset{\sim}{S}\underset{\sim}{Z}\underset{\sim}{D} \quad \text{and} \quad \underset{\sim}{D}\underset{\sim}{T} = (\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{S}\underset{\sim}{Z})^{-1} \tag{3.8}$$

when $\underset{\sim}{D}^{-1}$ exists; and

$$\underset{\sim}{D}\underset{\sim}{T} = \underset{\sim}{T}'\underset{\sim}{D} . \tag{3.9}$$

Then from (3.7) and (2.13)

$$\underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{T}^*\underset{\sim}{Z}'\underset{\sim}{R}^{-1} = \underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{Z}(\underset{\sim}{D}^{-1} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1} = \underset{\sim}{V}^{-1} . \tag{3.10}$$

Therefore

$$\begin{aligned}\underset{\sim}{Z}'\underset{\sim}{V}^{-1} &= \underset{\sim}{Z}'(\underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{T}^*\underset{\sim}{Z}'\underset{\sim}{R}^{-1}) \\ &= \underset{\sim}{Z}'\underset{\sim}{R}^{-1} - (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D} - \underset{\sim}{I})\underset{\sim}{T}^*\underset{\sim}{Z}'\underset{\sim}{R}^{-1} \\ &= \underset{\sim}{T}^*\underset{\sim}{Z}'\underset{\sim}{R}^{-1}, \quad \text{using (3.6)}, &(3.11) \\ &= (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1} . &(3.12)\end{aligned}$$

Similarly, from (2.46)

$$\underset{\sim}{Z}'\underset{\sim}{P} = \underset{\sim}{Z}'\underset{\sim}{S} - \underset{\sim}{Z}'\underset{\sim}{SZD}(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{SZD})^{-1}\underset{\sim}{Z}'\underset{\sim}{S}$$

$$= [\underset{\sim}{I} - (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{SZD} - \underset{\sim}{I})(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{SZD})^{-1}]\underset{\sim}{Z}'\underset{\sim}{S}$$

$$= (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{SZD})^{-1}\underset{\sim}{Z}'\underset{\sim}{S} \tag{3.13}$$

$$= \underset{\sim}{TZ}'\underset{\sim}{S} . \tag{3.14}$$

3.3. SPECIFIC SOLUTIONS

A specific solution of consistent equations $\underset{\sim}{Ax} = \underset{\sim}{\beta}$ is taken as being $\underset{\sim}{\tilde{x}} = \underset{\sim}{A}^{-}\underset{\sim}{\beta}$ for any given generalized inverse $\underset{\sim}{A}^{-}$ of $\underset{\sim}{A}$; and for an arbitrary vector $\underset{\sim}{z}$, general solutions shall mean $\underset{\sim}{\tilde{x}}^{o} = \underset{\sim}{A}^{-}\underset{\sim}{\beta} + (\underset{\sim}{A}^{-}\underset{\sim}{A} - \underset{\sim}{I})\underset{\sim}{z} = \underset{\sim}{\tilde{x}} + (\underset{\sim}{A}^{-}\underset{\sim}{A} - \underset{\sim}{I})\underset{\sim}{z}$, an expression that generates all possible solutions for any given $\underset{\sim}{A}^{-}$ (Theorem 3, Searle [1971a, p. 11]). For example, a specific solution of (3.1) is

$$\underset{\sim}{\hat{\alpha}} = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} \tag{3.15}$$

and general solutions are

$$\underset{\sim}{\hat{\alpha}}^{o} = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} + [(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} - \underset{\sim}{I}]\underset{\sim}{z}$$

$$= \underset{\sim}{\hat{\alpha}} + [(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} - \underset{\sim}{I}]\underset{\sim}{z} .$$

We now solve (3.3) and (3.4), observing first that the partitioned matrices on their left-hand sides are $\underset{\sim}{B}$ and $\underset{\sim}{C}$ defined respectively in (2.61) and (2.62). Using generalized inverses $\underset{\sim}{B}^{-}$ and $\underset{\sim}{C}^{-}$ of (2.65) and (2.63) respectively, and then $\underset{\sim}{B}^{\sim}$ and $\underset{\sim}{C}^{\sim}$ of (2.67) and (2.64), we obtain both specific and general solutions, in the sense just described, of (3.3) and (3.4). The reader of the sequel may reasonably feel that the presentation is excessive, and that, for example, it is necessary to develop only a specific solution to (3.3) using $\underset{\sim}{B}^{-}$, and use it to solve (3.4). Having developed a solution to each set of equations, it would then seem redundant to consider further solutions, obtained by other algebra. In a practical sense this is true. But since other solutions using algebra alternative to just B^- can be developed, and because some of them do not at first sight appear to be what might be expected, it seems that this Notebook is an appropriate place to show all details, tedious though they may be. It will surely be helpful to a thorough understanding of equations (3.3) and (3.4) to have available a compendium-like description of the different forms of solutions that can be developed. At the very least, it may save a reader or two from "re-inventing the wheel".

a. <u>Specific solutions using B^- and C^-</u>

With B^- of (2.65), a specific solution to (3.3) is

$$\begin{bmatrix} \tilde{\alpha} \\ \tilde{b} \end{bmatrix} = B^- \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \tag{3.16}$$

$$= \left\{ \begin{bmatrix} 0 & 0 \\ 0 & DT^* \end{bmatrix} + \begin{bmatrix} I \\ -DT^*Z'R^{-1}X \end{bmatrix} (X'V^{-1}X)^- [I \quad -X'R^{-1}ZDT^*] \right\} \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ DT^*Z'R^{-1}y \end{bmatrix} + \begin{bmatrix} I \\ -DT^*Z'R^{-1}X \end{bmatrix} (X'V^{-1}X)^- X'(I - R^{-1}ZDT^*Z')R^{-1}y$$

$$= \begin{bmatrix} 0 \\ DZ'V^{-1}y \end{bmatrix} + \begin{bmatrix} I \\ -DZ'V^{-1}X \end{bmatrix} (X'V^{-1}X)^- X'V^{-1}y, \qquad \text{from (3.11)},$$

$$= \begin{bmatrix} (X'V^{-1}X)^- X'V^{-1}y \\ DZ'V^{-1}[y - (X'V^{-1}X)^- X'V^{-1}y] \end{bmatrix}$$

$$= \begin{bmatrix} \hat{\alpha} \\ DZ'V^{-1}(y - X\hat{\alpha}) \end{bmatrix}, \qquad \text{from (3.15).} \tag{3.17}$$

Three features of (3.17) merit comment. First, a solution $\tilde{\alpha}$ in the MME's is the same as a solution $\hat{\alpha}$ of (3.15), in the generalized least squares equations. This is useful because on many occasions the MME's (3.3) are easier to solve computationally than (3.1), one reason being that (3.3) involves matrix inverses of smaller order than does (3.1). Second, in (3.17)

$$\tilde{b} = DZ'V^{-1}(y - X\hat{\alpha}) , \tag{3.18}$$

and even though $\hat{\alpha}$ of (3.15) is not invariant to whatever generalized inverse $(X'V^{-1}X)^-$ is used in (3.15), the occurrence of $\hat{\alpha}$ in (3.18) is such that $\tilde{b}$ <u>is</u> invariant to $(X'V^{-1}X)^-$; i.e., $\tilde{b}$ is the same for all solutions $\hat{\alpha}$. Third, under normality assumptions (which are often made), $\tilde{b}$ is exactly the same as the conditional expected value $E(b|y)$ save for $\hat{\alpha}$ in place of α. Since, under normality, $\hat{\alpha}$ is the ML estimator of α when D and V are known, we can then refer to

$$\tilde{b} = \widehat{E(b|y)} = DZ'V^{-1}(y - X\hat{\alpha}) \tag{3.19}$$

as the ML estimator of the conditional mean $E(b|y)$. It is in this sense that (3.19) is used in animal breeding programs where $\tilde{b}$ is often referred to as the BLUP (<u>b</u>est <u>l</u>inear <u>u</u>nbiased <u>p</u>redictor) of b.

Solution of (3.4) proceeds in a similar manner using C^- of (2.63) and without invoking D^{-1} (in case it does not exist):

$$\begin{bmatrix} \tilde{\alpha} \\ \tilde{\nu} \end{bmatrix} = C^- \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \tag{3.20}$$

$$= \left\{ \begin{bmatrix} 0 & 0 \\ 0 & T^* \end{bmatrix} + \begin{bmatrix} I \\ -T^*Z'R^{-1}X \end{bmatrix} (X'V^{-1}X)^-[I \quad -X'R^{-1}ZDT^*] \right\} \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ T^*Z'R^{-1}y \end{bmatrix} + \begin{bmatrix} I \\ -T^*Z'R^{-1}X \end{bmatrix} (X'V^{-1}X)^-[R^{-1} - R^{-1}ZD(I + Z'R^{-1}ZD)^{-1}Z'R^{-1}]y ,$$

after using (3.6); and then with the aid of (2.14) and (3.11) this becomes

$$\begin{bmatrix} \tilde{\boldsymbol{\alpha}} \\ \tilde{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\alpha}}) \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\alpha}} \\ \mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}}) \end{bmatrix} . \tag{3.21}$$

Comparison of (3.21) with (3.17) reveals that solutions $\tilde{\boldsymbol{\alpha}}$ are the same for (3.4) as for (3.3), namely $\tilde{\boldsymbol{\alpha}} = \hat{\boldsymbol{\alpha}}$; and from (3.21), $\mathbf{D}\tilde{\boldsymbol{\nu}} = \tilde{\mathbf{b}}$ of (3.19), in keeping with (3.5).

b. <u>Useful equalities</u>

Although the solution $\hat{\boldsymbol{\alpha}}$ in (3.15) may not be as convenient a way of computing $\hat{\boldsymbol{\alpha}}$ as is $\hat{\boldsymbol{\alpha}} = \tilde{\boldsymbol{\alpha}}$ from (3.3) or (3.4), a variety of useful equalities are based on $\hat{\boldsymbol{\alpha}}$ of (3.15). First,

$$\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}]\mathbf{y} = \mathbf{VPy} . \tag{3.22}$$

Then

$$\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}}) = \mathbf{Py} \tag{3.23}$$

$$= [\mathbf{S} - \mathbf{SZD}(\mathbf{I} + \mathbf{Z}'\mathbf{SZD})^{-1}\mathbf{Z}'\mathbf{S}]\mathbf{y}, \quad \text{from (2.46)} \tag{3.24}$$

and from (3.21)

$$\tilde{\boldsymbol{\nu}} = \mathbf{Z}'\mathbf{Py} \tag{3.25}$$

$$= (\mathbf{I} + \mathbf{Z}'\mathbf{SZD})^{-1}\mathbf{Z}'\mathbf{Sy}, \quad \text{from (3.13).} \tag{3.26}$$

Hence from (3.23) and (3.26)

$$\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}}) = \mathbf{S}(\mathbf{y} - \mathbf{ZD}\tilde{\boldsymbol{\nu}}) , \tag{3.27}$$

$$= \mathbf{S}(\mathbf{y} - \mathbf{Z}\tilde{\mathbf{b}}), \quad \text{from (3.5)} \tag{3.28}$$

$$= \mathbf{S}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\alpha}} - \mathbf{Z}\tilde{\mathbf{b}}), \quad \because \mathbf{SX} = \mathbf{0}, \tag{3.29}$$

$$= [\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{R}^{-1}](\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\alpha}} - \mathbf{Z}\tilde{\mathbf{b}}), \quad \text{from (2.36)}$$

$$= \mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\alpha}} - \mathbf{Z}\tilde{\mathbf{b}}), \tag{3.30}$$

on using the first equation in (3.3). Finally,

$$(\underset{\sim}{y} - \underset{\sim\sim}{X\hat{\alpha}})'\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim\sim}{X\hat{\alpha}}) = (\underset{\sim}{y}' - \underset{\sim}{\hat{\alpha}}'\underset{\sim}{X}')\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim\sim}{X\hat{\alpha}})$$

$$= \underset{\sim}{y}'\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim\sim}{X\hat{\alpha}}), \quad \because \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} = \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim\sim}{X\hat{\alpha}} \text{ as in (3.1)}, \tag{3.31}$$

$$= \underset{\sim}{y}'\underset{\sim\sim}{Py}, \quad \text{from (3.23)} \tag{3.32}$$

$$= (\underset{\sim}{y}' - \underset{\sim}{\hat{\alpha}}'\underset{\sim}{X}')\underset{\sim}{S}(\underset{\sim}{y} - \underset{\sim\sim}{Z\tilde{b}}), \quad \text{from (3.23)}$$

$$= \underset{\sim}{y}'\underset{\sim}{S}(\underset{\sim}{y} - \underset{\sim\sim}{Z\tilde{b}}), \quad \because \underset{\sim}{X}'\underset{\sim}{S} = \underset{\sim}{0}, \tag{3.33}$$

$$= (\underset{\sim}{y}' - \underset{\sim}{\hat{\alpha}}'\underset{\sim}{X}')\underset{\sim}{R}^{-1}(\underset{\sim}{y} - \underset{\sim\sim}{X\hat{\alpha}} - \underset{\sim\sim}{Z\tilde{b}}), \quad \text{from (3.30)}$$

$$= \underset{\sim}{y}'\underset{\sim}{R}^{-1}(\underset{\sim}{y} - \underset{\sim\sim}{X\hat{\alpha}} - \underset{\sim\sim}{Z\tilde{b}}), \tag{3.34}$$

after using the first equation in (3.3).

A number of the preceding results are to be found in Harville [1977]. The cross-references are as follows:

Cross-References to Harville [1977]

Equation No.	Harville	Equation No.	Harville
(3.1)	(3.1)	(3.12)	Following (3.6)
(3.3)	(3.8)	(3.13)	(3.7)
(3.4)	(3.3)	(3.21), $\tilde{v}$	Following (3.1)
(3.5)	Following (3.1)	(3.25)	Following (3.1)
(3.6)	Following (6.2)	(3.26)	(3.5)
(3.8)	Preceding (5.4)	(3.28) and (3.30)	(5.2)

c. Alternative solutions, using $B^{\sim}$ and $C^{\sim}$

Using $B^{\sim}$ of (2.67), a solution to (3.3) is

$$\begin{bmatrix} \tilde{\alpha} \\ \tilde{b} \end{bmatrix} = B^{\sim} \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

$$= \begin{bmatrix} (X'R^{-1}X)^{-}X'R^{-1}y \\ 0 \end{bmatrix} + \begin{bmatrix} -(X'R^{-1}X)^{-}X'R^{-1}Z \\ I \end{bmatrix} DTZ'[R^{-1} - R^{-1}X(X'R^{-1}X)^{-}X'R^{-1}]y \qquad (3.35)$$

$$= \begin{bmatrix} (X'R^{-1}X)^{-}X'R^{-1}(I - ZDTZ'S)y \\ DTZ'Sy \end{bmatrix}, \quad \text{using (2.36)}$$

$$= \begin{bmatrix} (X'R^{-1}X)^{-}X'R^{-1}(I - ZDZ'P)y \\ DZ'Py \end{bmatrix}, \quad \text{from (3.14)}$$

$$= \begin{bmatrix} (X'R^{-1}X)^{-}X'R^{-1}[I - (V - R)P]y \\ DZ'V^{-1}(y - X\hat{\alpha}) \end{bmatrix}, \quad \text{from (1.12) and (3.23)}$$

$$= \begin{bmatrix} (X'R^{-1}X)^{-}X'R^{-1}(y - VPy) \\ DZ'V^{-1}(y - X\hat{\alpha}) \end{bmatrix}, \quad \text{because } X'P = 0$$

$$= \begin{bmatrix} (X'R^{-1}X)^{-}X'R^{-1}X\hat{\alpha} \\ DZ'V^{-1}(y - X\hat{\alpha}) \end{bmatrix}, \quad \text{from (3.22).} \qquad (3.36)$$

Compared to (3.17) derived from B^{-}, the solution (3.36) derived from $B^{\sim}$ appears to have a different solution for $\tilde{\alpha}$. But in fact it does not: from (3.36) and (3.15)

$$\tilde{\underset{\sim}{\alpha}} = (\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim\sim}{X\hat{\alpha}} = (\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} \; . \tag{3.37}$$

Now consider

$$\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}[(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}]\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}$$

$$= \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{R}[\underset{\sim}{R}^{-1}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X}](\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}$$

$$= \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{R}[\underset{\sim}{R}^{-1}\underset{\sim}{X}](\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}, \quad \text{on applying (iv) of (2.9)}$$

$$= \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}$$

$$= \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} \; . \tag{3.38}$$

The matrix enclosed by [ ] on the left-hand side is the matrix that pre-multiplies $\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y}$ in (3.37); and (3.38) shows that it is a generalized inverse of $\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}$. Therefore (3.37) can be expressed as $\tilde{\underset{\sim}{\alpha}} = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y}$; hence $\tilde{\underset{\sim}{\alpha}}$ is a value of $\hat{\underset{\sim}{\alpha}}$.

Similarly, using $\underset{\sim}{C}^{\sim}$ of (2.64), a solution to (3.4) is

$$\begin{bmatrix} \tilde{\underset{\sim}{\alpha}} \\ \tilde{\underset{\sim}{\nu}} \end{bmatrix} = \underset{\sim}{C}^{\sim}\begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{y} \\ \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{y} \end{bmatrix}$$

$$= \begin{bmatrix} (\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{y} \\ \underset{\sim}{0} \end{bmatrix} + \begin{bmatrix} -(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim\sim}{ZD} \\ \underset{\sim}{I} \end{bmatrix}\underset{\sim\sim}{TZ}'[\underset{\sim}{R}^{-1} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}]\underset{\sim}{y} \tag{3.39}$$

and on comparing this with (3.35) and its ultimate form (3.36) it is clear that (3.37) becomes

$$\begin{bmatrix} \tilde{\underset{\sim}{\alpha}} \\ \tilde{\underset{\sim}{\nu}} \end{bmatrix} = \begin{bmatrix} (\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim\sim}{X\hat{\alpha}} \\ \underset{\sim}{Z}'\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim\sim}{X\hat{\alpha}}) \end{bmatrix}, \tag{3.40}$$

the same as for $\tilde{\underset{\sim}{\alpha}}$ in (3.36); and with $\underset{\sim\sim}{D\tilde{\nu}} = \tilde{\underset{\sim}{b}}$ as with (3.21) and (3.17).

3.4. GENERAL SOLUTIONS

a. Using $\mathbf{B}^-$ and $\mathbf{C}^-$

General solutions to (3.3) corresponding to the specific solutions are

$$\begin{bmatrix} \tilde{\boldsymbol{\alpha}}^o \\ \tilde{\mathbf{b}}^o \end{bmatrix} = \mathbf{B}^- \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} + (\mathbf{B}^-\mathbf{B} - \mathbf{I})\mathbf{z} \tag{3.41}$$

where $\mathbf{z}$ is arbitrary, of appropriate order. The first part of (3.41) is given in (3.17) as a specific solution; the second part requires $\mathbf{B}^-\mathbf{B}$ which, from (2.65) and (2.61), is

$$\mathbf{B}^-\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{DT}^{\#}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{DT}^{\#}(\mathbf{D}^{-1} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}) \end{bmatrix}$$

$$+ \begin{bmatrix} \mathbf{I} \\ -\mathbf{DT}^{\#}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} \end{bmatrix} (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^-[\mathbf{X}'(\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{ZDT}^{\#}\mathbf{Z}'\mathbf{R}^{-1})\mathbf{X} \;\vdots\; \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} - \mathbf{X}'\mathbf{R}^{-1}\mathbf{ZDT}^{\#}(\mathbf{D}^{-1} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z})]$$

and on using (3.6), (3.10) and (3.11) this becomes

$$\mathbf{B}^-\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{DZ}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{I} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -\mathbf{DZ}'\mathbf{V}^{-1}\mathbf{X} \end{bmatrix} (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^-[\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} \quad \mathbf{0}]$$

$$= \begin{bmatrix} (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad \because \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} = \mathbf{X}\,. \tag{3.42}$$

Substituting this and (3.17) into (3.41) gives general solutions to (3.3) using $\mathbf{B}^-$ as

$$\begin{bmatrix} \tilde{\underset{\sim}{\alpha}}^{o} \\ \tilde{\underset{\sim}{b}}^{o} \end{bmatrix} = \begin{bmatrix} \hat{\underset{\sim}{\alpha}} + [(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} - \underset{\sim}{I}]\underset{\sim}{z} \\ \underset{\sim}{D}\underset{\sim}{Z}'\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\hat{\underset{\sim}{\alpha}}) \end{bmatrix} = \begin{bmatrix} \hat{\underset{\sim}{\alpha}}^{o} \\ \tilde{\underset{\sim}{b}} \end{bmatrix} \tag{3.43}$$

for any arbitrary $\underset{\sim}{z}$ of order p.

This result is interesting, because it shows that general solutions for $\tilde{\underset{\sim}{\alpha}}^{o}$ correspond precisely to those for $\hat{\underset{\sim}{\alpha}}^{o}$ following (3.15); and, of course, there is only the one solution $\tilde{\underset{\sim}{b}}$.

In the same way that $(\underset{\sim}{B}^{-}\underset{\sim}{B} - \underset{\sim}{I})\underset{\sim}{z}$ in (3.41) is added to the specific solution of (3.16) to obtain general solutions of (3.3) using $\underset{\sim}{B}^{-}$, so also will $(\underset{\sim}{C}^{-}\underset{\sim}{C} - \underset{\sim}{I})\underset{\sim}{z}$ be added to (3.20) to yield general solutions to (3.4) using $\underset{\sim}{C}^{-}$. For this, from (2.63) and (2.62)

$$\underset{\sim}{C}^{-}\underset{\sim}{C} = \begin{bmatrix} \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{T}^{*}\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{T}^{*}(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}) \end{bmatrix}$$

$$+ \begin{bmatrix} \underset{\sim}{I} \\ -\underset{\sim}{T}^{*}\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{X} \end{bmatrix} (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{Z})^{-}[\underset{\sim}{X}'(\underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{T}^{*}\underset{\sim}{Z}'\underset{\sim}{R}^{-1})\underset{\sim}{X} \;\vdots\; \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D} - \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{T}^{*}(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D})]$$

and on using (3.6), (2.14) and (3.11) this becomes

$$\underset{\sim}{C}^{-}\underset{\sim}{C} = \begin{bmatrix} \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{Z}'\underset{\sim}{V}^{-1}\underset{\sim}{X} & \underset{\sim}{I} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{I} \\ -\underset{\sim}{Z}'\underset{\sim}{V}^{-1}\underset{\sim}{X} \end{bmatrix} (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}[\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} \quad \underset{\sim}{0}]$$

$$= \begin{bmatrix} (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{I} \end{bmatrix} .$$

Adding $(\underset{\sim}{C}^{-}\underset{\sim}{C} - \underset{\sim}{I})\underset{\sim}{z}$ to (3.20) therefore yields general solutions of (3.4) as

$$\begin{bmatrix}\tilde{\alpha}^o\\ \tilde{\nu}^o\end{bmatrix} = \begin{bmatrix}\hat{\alpha} + [(X'V^{-1}X)^- X'V^{-1}X - I]z \\ Z'V^{-1}(y - X\hat{\alpha})\end{bmatrix} \qquad (3.44)$$

in agreement with (3.43): $\hat{\alpha}^o$ is the same, and $D\tilde{\nu} = \tilde{b}$, just as in the comparison of specific solutions (3.21) and (3.17).

b. Using $B^{\sim}$ and $C^{\sim}$

In a similar vein, to convert specific solutions (3.36) using $B^{\sim}$, we derive $B^{\sim}B$: from (2.67) and (2.61) it is

$$B^{\sim}B = \begin{bmatrix}(X'R^{-1}X)^-X'R^{-1}X & (X'R^{-1}X)^-X'R^{-1}Z \\ 0 & 0\end{bmatrix}$$

$$+ \begin{bmatrix}-(X'R^{-1}X)^-X'R^{-1}Z \\ I\end{bmatrix} DT\{Z'[R^{-1} - R^{-1}X(X'R^{-1}X)^-X'R^{-1}]X \;\vdots\; D^{-1} + Z'[R^{-1} - R^{-1}X(X'R^{-1}X)^-X'R^{-1}]Z\}.$$

Using S from (2.36) this is

$$B^{\sim}B = \begin{bmatrix}(X'R^{-1}X)^-X'R^{-1}X & (X'R^{-1}X)^-X'R^{-1}Z \\ 0 & 0\end{bmatrix} + \begin{bmatrix}-(X'R^{-1}X)^-X'R^{-1}Z \\ I\end{bmatrix}[DTZ'SX \quad DT(D^{-1} + Z'SZ)]$$

and with $DTZ'SX = DZ'PX$ from (3.14) and $PX = 0$ from (2.27), together with (3.8), this becomes

$$B^{\sim}B = \begin{bmatrix}(X'R^{-1}X)^-X'R^{-1}X & (X'R^{-1}X)^-X'R^{-1}Z \\ 0 & 0\end{bmatrix} + \begin{bmatrix}-(X'R^{-1}X)^-X'R^{-1}Z \\ I\end{bmatrix}[0 \quad I]$$

$$= \begin{bmatrix}(X'R^{-1}X)^-X'R^{-1}X & 0 \\ 0 & I\end{bmatrix}.$$

Hence adding $(\underset{\sim}{B}^{\sim}\underset{\sim}{B} - \underset{\sim}{I})\underset{\sim}{w}$ to (3.36) gives general solutions to (3.3) using $\underset{\sim}{B}^{\sim}$ as

$$\begin{bmatrix} \underset{\sim}{\tilde{\alpha}}^{o} \\ \underset{\sim}{\tilde{b}}^{o} \end{bmatrix} = \begin{bmatrix} (\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X}(\underset{\sim}{\hat{\alpha}} + \underset{\sim}{w}) - \underset{\sim}{w} \\ \underset{\sim}{D}\underset{\sim}{Z}'\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\hat{\alpha}}) \end{bmatrix} \tag{3.45}$$

for $\underset{\sim}{w}$ arbitrary and, in (3.45), of order p.

Comparing (3.45) with (3.43), the general solutions to (3.3) using $\underset{\sim}{B}^{-}$, we see that $\underset{\sim}{\tilde{b}}$ is, naturally, the same in both places — but $\underset{\sim}{\tilde{\alpha}}^{o}$ does not appear to be. Nevertheless, the $\underset{\sim}{\tilde{\alpha}}^{o}$'s of (3.43) and (3.45) are equivalent, as we now show. In (3.45) write

$$\underset{\sim}{\tilde{\alpha}}^{o} = \underset{\sim}{\hat{\alpha}} + [(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X} - \underset{\sim}{I}](\underset{\sim}{\hat{\alpha}} + \underset{\sim}{w}) \ . \tag{3.46}$$

Then (3.46) and $\underset{\sim}{\tilde{\alpha}}^{o}$ of (3.43) will be the same if $\underset{\sim}{w}$ is such that

$$[(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X} - \underset{\sim}{I}](\underset{\sim}{\hat{\alpha}} + \underset{\sim}{w}) = [(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} - \underset{\sim}{I}]\underset{\sim}{z} \ . \tag{3.47}$$

Because $(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X} - \underset{\sim}{I}$ is idempotent, it is its own generalized inverse, and so a solution to (3.47) for $\underset{\sim}{\hat{\alpha}} + \underset{\sim}{w}$ is

$$\underset{\sim}{\hat{\alpha}} + \underset{\sim}{w} = [(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X} - \underset{\sim}{I}][(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} - \underset{\sim}{I}]\underset{\sim}{z} \ . \tag{3.48}$$

The equality $\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} = \underset{\sim}{X}$, an extension of (2.9), reduces (3.48) to

$$\underset{\sim}{\hat{\alpha}} + \underset{\sim}{w} = -[(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} - \underset{\sim}{I}]\underset{\sim}{z} \tag{3.49}$$

so leading to (3.46) being

$$\underset{\sim}{\tilde{\alpha}}^{o} = \underset{\sim}{\hat{\alpha}} + [(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} - \underset{\sim}{I}]\underset{\sim}{z}, \ = \underset{\sim}{\hat{\alpha}}^{o}$$

as in (3.41); i.e., for any $\underset{\sim}{w}$ satisfying (3.49) for arbitrary $\underset{\sim}{z}$, the general solutions $\underset{\sim}{\tilde{\alpha}}^{o}$ of (3.45), based on $\underset{\sim}{B}^{\sim}$, are the same as solutions $\underset{\sim}{\tilde{\alpha}}^{o} = \underset{\sim}{\hat{\alpha}}^{o}$ of (3.43) for that same $\underset{\sim}{z}$, based on $\underset{\sim}{B}^{-}$.

Finally consider $C^{-}C$; from (2.64) and (2.62) its value is

$$C^{-}C = \begin{bmatrix} (X'R^{-1}X)^{-}X'R^{-1}X & (X'R^{-1}X)^{-}X'R^{-1}ZD \\ 0 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} -(X'R^{-1}X)^{-}X'R^{-1}ZD \\ I \end{bmatrix} T\{Z'[R^{-1} - R^{-1}X(X'R^{-1}X)^{-}Z'R^{-1}]X \;\vdots\; I + Z'[R^{-1} - R^{-1}X(X'R^{-1}X)^{-}X'R^{-1}]ZD\}$$

$$= \begin{bmatrix} (X'R^{-1}X)^{-}X'R^{-1}X & (X'R^{-1}X)^{-}X'R^{-1}ZD \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -(X'R^{-1}X)^{-}X'R^{-1}ZD \\ I \end{bmatrix} [TZ'SX \quad T(I + Z'SZD)] .$$

Using (3.14) along with $PX = 0$ gives $TZ'SX = Z'PX = 0$, and (3.8) gives $T(I + Z'SZD) = I$ so that

$$C^{-}C = \begin{bmatrix} (X'R^{-1}X)^{-}X'R^{-1}X & 0 \\ 0 & I \end{bmatrix} . \tag{3.50}$$

Adding $(C^{-}C - I)w$ to the specific solution (3.40) therefore gives exactly the same result as (3.45) except for having $\tilde{v} = Z'V^{-1}(y - X\hat{\alpha})$ in place of $\tilde{b} = D\tilde{v}$, as one would expect.

Summary of Solutions

The results of practical importance are shown in Table 1, where $\hat{\alpha}$ is a value of

$$\hat{\alpha} = (X'V^{-1}X)^{-}X'V^{-1}y .$$

Table 1. Usable Solutions of MME's

Parameter Vector	Equation (3.3) (Henderson; requires $\mathbf{D}^{-1}$)	Equation (3.5) (Harville; $\mathbf{D}$ can be singular)
	Solutions	
$\boldsymbol{\alpha}$	$\hat{\boldsymbol{\alpha}}$	$\hat{\boldsymbol{\alpha}}$
$\mathbf{b}$	$\tilde{\mathbf{b}} = \mathbf{D}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}})$	$\tilde{\boldsymbol{\nu}} = \mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}})$ with $\tilde{\mathbf{b}} = \mathbf{D}\tilde{\boldsymbol{\nu}}$

Table 2 shows all the specific and general solutions using the notation

$$\hat{\boldsymbol{\alpha}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$$

$$\hat{\boldsymbol{\alpha}}^{o} = \hat{\boldsymbol{\alpha}} + [(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} - \mathbf{I}]\mathbf{z}, \quad \text{for arbitrary } \mathbf{z}$$

$$\tilde{\mathbf{b}} = \mathbf{D}^{-1}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}}), \quad \text{for non-singular } \mathbf{D}$$

$$\tilde{\boldsymbol{\nu}} = \mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}}), \quad \text{with } \tilde{\mathbf{b}} = \mathbf{D}\tilde{\boldsymbol{\nu}} \text{ for any } \mathbf{D} .$$

Table 2. Specific and General Solutions

Equations to be solved			Solutions			
			Generalized Inverses Used			
			$\underset{\sim}{B}^{-}$ and $\underset{\sim}{C}^{-}$		$\underset{\sim}{B}^{\sim}$ and $\underset{\sim}{C}^{\sim}$	
Equations to be solved	Characteristics	Parameter Vector	Equ.	Solution	Equ.	Solution
						Specific Solutions
(3.3)	Matrix $\underset{\sim}{B}$ (requires $\underset{\sim}{D}^{-1}$)	$\underset{\sim}{\alpha}$	(3.17)	$\underset{\sim}{\hat{\alpha}}$	(3.36)	$(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X}\underset{\sim}{\hat{\alpha}}$ = an $\underset{\sim}{\hat{\alpha}}$
		$\underset{\sim}{b}$		$\underset{\sim}{\tilde{b}}$		$\underset{\sim}{\tilde{b}}$
(3.4)	Matrix $\underset{\sim}{C}$ ($\underset{\sim}{D}$ can be singular)	$\underset{\sim}{\alpha}$	(3.21)	$\underset{\sim}{\hat{\alpha}}$	(3.40)	Same as (3.36)
		$\underset{\sim}{\nu}$		$\underset{\sim}{\tilde{\nu}}$		$\underset{\sim}{\tilde{\nu}}$
						General Solutions*
(3.3)	Matrix $\underset{\sim}{B}$ (requires $\underset{\sim}{D}^{-1}$)	$\underset{\sim}{\alpha}$	(3.43)	$\underset{\sim}{\hat{\alpha}}^{o}$	(3.45)	$(\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X}(\underset{\sim}{\hat{\alpha}}+\underset{\sim}{w})-\underset{\sim}{w}$ = an $\underset{\sim}{\hat{\alpha}}^{o}$
		$\underset{\sim}{b}$		$\underset{\sim}{\tilde{b}}$		$\underset{\sim}{\tilde{b}}$
(3.4)	Matrix $\underset{\sim}{C}$ ($\underset{\sim}{D}$ can be singular)	$\underset{\sim}{\alpha}$	(3.44)	$\underset{\sim}{\hat{\alpha}}^{o}$	(3.50)	Same as (3.45)
		$\underset{\sim}{\nu}$		$\underset{\sim}{\tilde{\nu}}$		$\underset{\sim}{\tilde{\nu}}$

* $\underset{\sim}{w} = -\underset{\sim}{\hat{\alpha}} - [(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} - \underset{\sim}{I}]\underset{\sim}{z}$, for arbitrary $\underset{\sim}{z}$.

Chapter 4

MAXIMUM LIKELIHOOD

The maximum likelihood (ML) method of estimating parameters is a solidly entrenched procedure in statistics. And rightly so. It has many optimum properties. In brief, it is a method which yields, as estimates of the parameters, values which, if they are used in place of the parameters in the likelihood function of the data (assumed to be a random sample from the population), yields a maximum value of that function. In a sense, therefore, ML estimators are values which, if they were the true parameter values would, on the basis of the chosen probability density function, yield a larger value for the likelihood of the observed sample than any other values of the parameters; i.e., the ML estimators maximize the likelihood.

Application of the ML method therefore entails assuming a probability density function for the random variables being studied, and then writing down the likelihood function of the sample of data. This is treated as a function of arguments represented by the parameters, and that function is then maximized with respect to those arguments. This usually involves differentiating the likelihood function (or its natural logarithm) with respect to the parameters, equating the results to zero and solving the resultant equations. The solutions must be constrained to lie in the parameter space. In the case of variance components estimation, these constraining conditions are $\sigma_i^2 \geq 0$ for $i = 1, 2, \cdots, c$ and $\sigma_0^2 > 0$. The ML estimators must satisfy similar conditions: $\tilde{\sigma}_i^2 \geq 0$ for $i = 1, 2, \cdots, c$ and $\tilde{\sigma}_0^2 > 0$.

Associated with the vector of ML estimators $\tilde{\boldsymbol{\sigma}}^2 = [\tilde{\sigma}_0^2 \;\; \tilde{\sigma}_1^2 \;\; \cdots \;\; \tilde{\sigma}_c^2]'$, corresponding to $\boldsymbol{\sigma}^2$ of (1.27), is a matrix called the information matrix. It is the basis for deriving sampling variances of, and covariances between, the ML estimators, i.e., for deriving var$(\tilde{\sigma}_0^2)$ and cov$(\tilde{\sigma}_0^2, \tilde{\sigma}_1^2)$ and other such terms.

The estimation equations and information matrices derived in this chapter are indicated in the following table.

Table 4.1. Equation Numbers in Chapter 4 of the Main Results for Maximum Likelihood (ML)

Parameter	Estimation Equations: Basic equations	Equations using MME's	Equations for $\tilde{\sigma}^2$
σ^2	(4.4) - (4.6), (4.16a)	(4.25) - (4.27)	(4.63)
γ	(4.11) - (4.13)		

Parameter	Information Matrices: Basic equations	Using MME's	Used in estimation equations with MME's	Used in Newton-Raphson equations
σ^2	(4.33)	(4.60)	(4.70)	(4.85) - (4.87)
γ	(4.37	(4.61)	(4.73)	(4.94)

4.1. DISTRIBUTIONAL ASSUMPTIONS

The ML procedure can be used for almost any probability density function, but for estimating variance components it is customary to assume normality. This means that we assume $\underset{\sim}{y}$ follows a multivariate normal distribution, which we denote by $\underset{\sim}{y} \sim \mathcal{N}(\underset{\sim}{X\alpha}, \underset{\sim}{V})$ where $E(\underset{\sim}{y}) = \underset{\sim}{X\alpha}$ and var$(\underset{\sim}{y}) = \underset{\sim}{V}$ as in (1.6) and (1.12). Then the logarithm of the likelihood function is

$$L = -\tfrac{1}{2}N \log 2\pi - \tfrac{1}{2} \log|\underset{\sim}{V}| - \tfrac{1}{2}(\underset{\sim}{y} - \underset{\sim}{X\alpha})'\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim}{X\alpha}) \ . \tag{4.1}$$

On using $\underset{\sim}{V} = \sigma_0^2\underset{\sim}{H}$ of (1.21) this can also be written as

$$L = -\tfrac{1}{2}N \log 2\pi - \tfrac{1}{2}N \log \sigma_0^2 - \tfrac{1}{2} \log|\underset{\sim}{H}| - (\underset{\sim}{y} - \underset{\sim}{X\alpha})'\underset{\sim}{H}^{-1}(\underset{\sim}{y} - \underset{\sim}{X\alpha})/2\sigma_0^2 \ . \tag{4.2}$$

4.2. TWO STANDARD RESULTS

We give two well-known results in the theory of maximum likelihood, stated in general terms. Suppose $f(\underset{\sim}{x}, \underset{\sim}{\theta})$ is the probability density function of a random variable $\underset{\sim}{X}$, satisfying the usual regularity conditions involving a vector of parameters $\underset{\sim}{\theta}' = [\theta_1 \cdots \theta_t]$. Then for a random sample $\underset{\sim}{x}' = [x_1 \cdots x_n]$ the logarithm of the likelihood is

$$L = \log f(\underset{\sim}{x}, \underset{\sim}{\theta}) = \log \prod_{i=1}^{n} f(x_i, \underset{\sim}{\theta}) = \sum_{i=1}^{n} \log f(x_i, \underset{\sim}{\theta}) .$$

The two results are stated as Lemmas.

Lemma 4.1. $$E\left(\frac{\partial L}{\partial \underset{\sim}{\theta}}\right) = 0 .$$

Proof: $$E\left(\frac{\partial L}{\partial \theta_i}\right) = \int_{\underset{\sim}{x}} \frac{\partial \log f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial \theta_i} f(\underset{\sim}{x}, \underset{\sim}{\theta}) d\underset{\sim}{x}$$

$$= \int_{\underset{\sim}{x}} \frac{1}{f(\underset{\sim}{x}, \underset{\sim}{\theta})} \frac{\partial f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial \theta_i} f(\underset{\sim}{x}, \underset{\sim}{\theta}) d\underset{\sim}{x}$$

$$= \int_{\underset{\sim}{x}} \frac{\partial}{\partial \theta_i} f(\underset{\sim}{x}, \underset{\sim}{\theta}) d\underset{\sim}{x} = \frac{\partial}{\partial \theta_i} \int_{\underset{\sim}{x}} f(\underset{\sim}{x}, \underset{\sim}{\theta}) d\underset{\sim}{x} = \frac{\partial}{\partial \theta_i} 1 = 0. \quad \text{Q.E.D.}$$

Lemma 4.2. $$E\left\{\frac{\partial^2 L}{\partial \theta_i \partial \theta_j}\right\} \text{ for } i,j = 1, \cdots, t, = -E\left(\frac{\partial L}{\partial \underset{\sim}{\theta}}\right)\left(\frac{\partial L}{\partial \underset{\sim}{\theta}}\right)' .$$

Proof: $$\text{LHS} = E\left\{\frac{\partial}{\partial \theta_i}\left(\frac{\partial \log f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial \theta_j}\right)\right\} = E\left\{\frac{\partial}{\partial \theta_i}\left(\frac{1}{f(\underset{\sim}{x}, \underset{\sim}{\theta})} \frac{\partial f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial \theta_j}\right)\right\}$$

$$= E\left\{\frac{-1}{[f(\underset{\sim}{x}, \underset{\sim}{\theta})]^2} \frac{\partial f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial \theta_i} \frac{\partial f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial \theta_j} + \frac{1}{f(\underset{\sim}{x}, \underset{\sim}{\theta})} \frac{\partial^2 f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial \theta_i \partial \theta_j}\right\}$$

$$= \left\{-\int \frac{\partial f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial\theta_i} \frac{\partial f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial\theta_j} \frac{1}{[f(\underset{\sim}{x}, \underset{\sim}{\theta})]^2} f(\underset{\sim}{x}, \underset{\sim}{\theta})d\underset{\sim}{x} + \int \frac{\partial^2 f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial\theta_i\partial\theta_j} \frac{f(\underset{\sim}{x}, \underset{\sim}{\theta})}{f(\underset{\sim}{x}, \underset{\sim}{\theta})} d\underset{\sim}{x}\right\}$$

$$= \left\{-\int \frac{1}{f(\underset{\sim}{x}, \underset{\sim}{\theta})} \frac{\partial f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial\theta_i} \frac{1}{f(\underset{\sim}{x}, \underset{\sim}{\theta})} \frac{\partial f(\underset{\sim}{x}, \underset{\sim}{\theta})}{\partial\theta_j} f(\underset{\sim}{x}, \underset{\sim}{\theta})d\underset{\sim}{x} + \frac{\partial^2}{\partial\theta_i\partial\theta_j} \int f(\underset{\sim}{x}, \underset{\sim}{\theta})d\underset{\sim}{x}\right\}$$

$$= \left\{-\int \frac{\partial}{\partial\theta_i} \log f(\underset{\sim}{x}, \underset{\sim}{\theta}) \frac{\partial}{\partial\theta_j} \log f(\underset{\sim}{x}, \underset{\sim}{\theta}) f(\underset{\sim}{x}, \underset{\sim}{\theta})d\underset{\sim}{x} + \frac{\partial^2(1)}{\partial\theta_i\partial\theta_j}\right\}$$

$$= -E \frac{\partial L}{\partial\underset{\sim}{\theta}} \left(\frac{\partial L}{\partial\underset{\sim}{\theta}}\right)' = \text{RHS} . \qquad \underline{\text{Q.E.D.}}$$

4.3. EQUATIONS TO BE SOLVED

a. For $\underset{\sim}{\alpha}$ and $\underset{\sim}{\sigma}^2$

Equations for obtaining $\tilde{\underset{\sim}{\alpha}}$ and $\tilde{\underset{\sim}{V}}$, the ML estimators of $\underset{\sim}{\alpha}$ and $\underset{\sim}{V}$, come from differentiating (4.1), making use of (2.1) and (2.4):

$$\frac{\partial L}{\partial\underset{\sim}{\alpha}} = -\tfrac{1}{2}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}\underset{\sim}{\alpha} - \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y})$$

$$\frac{\partial L}{\partial\sigma_i^2} = -\tfrac{1}{2} \operatorname{tr}\left(\underset{\sim}{V}^{-1} \frac{\partial\underset{\sim}{V}}{\partial\sigma_i^2}\right) + \tfrac{1}{2}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})'\underset{\sim}{V}^{-1} \frac{\partial\underset{\sim}{V}}{\partial\sigma_i^2} \underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha}) . \qquad (4.3)$$

Equating these to zero gives the equations

$$\underset{\sim}{X}'\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{X}\tilde{\underset{\sim}{\alpha}} = \underset{\sim}{X}'\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{y} , \qquad (4.4)$$

and

$$\operatorname{tr}(\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') = (\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\alpha}})'\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{V}}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\alpha}}) \qquad (4.5)$$

$$= \underset{\sim}{y}'\tilde{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{P}}\underset{\sim}{y}, \qquad \text{for } i = 0, 1, \cdots, c, \qquad (4.6)$$

after using $\partial\mathbf{V}/\partial\sigma_i^2 = \mathbf{Z}_i\mathbf{Z}_i'$ of (2.3), and (3.23).

The symbol $\tilde{\boldsymbol{\alpha}}$ here is <u>not</u> the same as $\tilde{\boldsymbol{\alpha}} \equiv \hat{\boldsymbol{\alpha}}$ occurring in the MME's of Chapter 3. In that situation $\tilde{\boldsymbol{\alpha}} = \hat{\boldsymbol{\alpha}}$ is, as we saw, a solution to the generalized least squares equations (3.1). But $\tilde{\boldsymbol{\alpha}}$ here is defined by (4.4), which, although it has the appearance of the generalized least squares equations (3.1), does in fact have the important difference that (4.4) has $\tilde{\mathbf{V}}$ in place of $\mathbf{V}$; i.e., $\tilde{\boldsymbol{\alpha}}$ is $\hat{\boldsymbol{\alpha}}$ using $\tilde{\mathbf{V}}$ in place of $\mathbf{V}$.

Equations (4.4) and (4.5) are to be solved for $\tilde{\boldsymbol{\alpha}}$ and for the $\tilde{\boldsymbol{\sigma}}^2$'s implicit in $\tilde{\mathbf{V}}$. The solutions are to be distinguished from the ML estimators. Only non-negative solutions $\tilde{\sigma}_i^2$ can be considered as ML estimators — because, by definition, $\sigma_0^2 > 0$ and $\sigma_i^2 \geq 0$, $i = 1, \cdots, c$, and ML estimators must satisfy these conditions. This non-negativity constraint is a difficulty that must be taken into account in computer programs that are used for solving equations (4.4) and (4.5). It is customary for any $\tilde{\sigma}_i^2$ that is computed as a negative value to be put equal to zero — an action which has the effect, of course, of altering the model being used. It also raises the further difficulty of having a computer program which, for any $\tilde{\sigma}_i^2$ that has been put equal to zero after some iteration, enables that $\tilde{\sigma}_i^2$ to come back into the calculations again at some later iteration if it were then to be positive. Computing difficulties of this nature are considered in such papers as Hemmerle and Hartley [1973] and Jennrich and Sampson [1976].

<u>b. For $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$</u>

The ML equations for $\boldsymbol{\alpha}$, σ_0^2 and $\boldsymbol{\gamma}$ come from differentiating (4.2):

$$\frac{\partial L}{\partial \boldsymbol{\alpha}} = -\tfrac{1}{2}(\mathbf{X}'\mathbf{H}^{-1}\mathbf{y} - \mathbf{X}'\mathbf{H}^{-1}\mathbf{X}\boldsymbol{\alpha})/\sigma_0^2 \tag{4.7}$$

$$\frac{\partial L}{\partial \gamma_0} = -\tfrac{1}{2}N/\sigma_0^2 + \tfrac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\alpha})'\mathbf{H}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\alpha})/\sigma_0^4 \tag{4.8}$$

$$= -\tfrac{1}{2}[N\sigma_0^2 - (\mathbf{y} - \mathbf{X}\boldsymbol{\alpha})'\mathbf{H}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\alpha})]/\sigma_0^4 \tag{4.9}$$

and

$$\frac{\partial L}{\partial \gamma_i} = -\tfrac{1}{2}\mathrm{tr}(\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') + \tfrac{1}{2}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})/\sigma_0^2 , \qquad (4.10)$$

$$\text{for } i = 1, 2, \cdots, c.$$

The ML equations are then

$$\underset{\sim}{X}'\underset{\sim}{\tilde{H}}^{-1}\underset{\sim}{X}\underset{\sim}{\tilde{\alpha}} = \underset{\sim}{X}'\underset{\sim}{\tilde{H}}^{-1}\underset{\sim}{y} \qquad (4.11)$$

$$\tilde{\sigma}_0^2 = (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}})'\underset{\sim}{\tilde{H}}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}})/N \qquad (4.12)$$

$$\mathrm{tr}(\underset{\sim}{\tilde{H}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') = (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}})'\underset{\sim}{\tilde{H}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{\tilde{H}}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}})/\tilde{\sigma}_0^2 \quad \text{for } i = 1, \cdots, c. \qquad (4.13)$$

These are the equations derived by Hartley and Rao [1967]. Solutions satisfying $\tilde{\sigma}_0^2 > 0$ and $\tilde{\gamma}_i \geq 0$ are ML estimators.

c. Equivalence of the 2 sets of equations

Equations (4.11), (4.12) and (4.13) are obviously equivalent to (4.4) and (4.5). But it is instructive to demonstrate this equivalence. Certainly, on using $\underset{\sim}{V}^{-1} = \underset{\sim}{H}^{-1}/\sigma_0^2$ from (1.21) it is clear that (4.4) and (4.11) are the same. Likewise for i = 1, $\cdots$, c, (4.6) and (4.13) are the same. It remains to show that (4.5) for i = 0 is the same as (4.12) — an equivalence which is certainly not obvious. Put i = 0 in (4.5) and use $\underset{\sim}{Z}_0 = \underset{\sim}{I}$ of (1.23) and get

$$\mathrm{tr}(\underset{\sim}{\tilde{V}}^{-1}) = (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}})'\underset{\sim}{\tilde{V}}^{-2}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}}) ,$$

i.e.,

$$\mathrm{tr}(\underset{\sim}{\tilde{H}}^{-1}) = (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}})'\underset{\sim}{\tilde{H}}^{-2}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}})/\tilde{\sigma}_0^2 \qquad (4.14)$$

or

$$\tilde{\sigma}_0^2 = \frac{(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}})'\underset{\sim}{\tilde{H}}^{-2}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{\alpha}})}{\mathrm{tr}(\underset{\sim}{\tilde{H}}^{-1})} . \qquad (4.15)$$

This is supposedly (4.12). We show that it is, by showing that (4.12) and (4.13) lead directly to (4.15). Multiply (4.13) by $\tilde{\sigma}_i^2$ and add over i = 1, ⋯, c to get

$$\mathrm{tr}(\tilde{H}^{-1} \sum_{i=1}^{c} Z_i Z_i' \tilde{\sigma}_i^2) = (y - X\tilde{\alpha})' \tilde{H}^{-1} (\sum_{i=1}^{c} Z_i Z_i' \tilde{\sigma}_i^2) \tilde{H}^{-1} (y - X\tilde{\alpha}) / \tilde{\sigma}_0^2 . \qquad (4.16)$$

Now use (1.22) and (1.20) to write

$$\tilde{H} - I = \sum_{i=1}^{c} Z_i Z_i' \tilde{\sigma}_i^2 / \tilde{\sigma}_0^2$$

and so (4.16) becomes

$$\mathrm{tr}[\tilde{H}^{-1}(\tilde{H} - I)\tilde{\sigma}_0^2] = (y - X\tilde{\alpha})' \tilde{H}^{-1} (\tilde{H} - I) \tilde{H}^{-1} (y - X\tilde{\alpha})$$

$$\tilde{\sigma}_0^2 [N - \mathrm{tr}(\tilde{H}^{-1})] = (y - X\tilde{\alpha})' (\tilde{H}^{-1} - \tilde{H}^{-2}) (y - X\tilde{\alpha}) ;$$

and subtracting this from (4.12) gives (4.14), which leads to (4.15). Hence the two sets of equations are equivalent. Therefore (4.12) and (4.13) can be written as

$$\mathrm{tr}(\tilde{H}^{-1} Z_i Z_i') = (y - X\tilde{\alpha})' \tilde{H}^{-1} Z_i Z_i' \tilde{H}^{-1} (y - X\tilde{\alpha}) / \tilde{\sigma}_0^2 \quad \text{for } i = 0, 1, \cdots, c. \qquad (4.16a)$$

d. Comparison with MME's

We show here how the ML equations (4.12) and (4.13) can be expressed in terms of portions of the MME's of Section 3. First, equation (4.12): it is

$$\begin{aligned} \tilde{\sigma}_0^2 &= (y - X\alpha)' \tilde{H}^{-1} (y - X\tilde{\alpha}) / N \\ &= \tilde{\sigma}_0^2 (y - X\tilde{\alpha})' \tilde{V}^{-1} (y - X\tilde{\alpha}) / N \\ &= \tilde{\sigma}_0^2 y' \tilde{R}^{-1} (y - X\tilde{\alpha} - Z\tilde{b}) / N \quad \text{from (3.34)} \\ &= y' (y - X\tilde{\alpha} - Z\tilde{b}) / N \quad \text{for } \tilde{R} = \tilde{\sigma}_0^2 I , \qquad (4.17) \end{aligned}$$

where $\tilde{\alpha}$ and $\tilde{b}$ are as defined in the MME's (3.3) or, equivalently (3.4) and (3.5) —

in either case, using the ML $\tilde{\mathbf{V}}$ in those equations in place of $\mathbf{V}$.

Equation (4.13) for $i = 1, \cdots, c$ is

$$\operatorname{tr}(\tilde{\mathbf{H}}^{-1}\mathbf{Z}_i\mathbf{Z}_i') = (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\alpha}})'\tilde{\mathbf{H}}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\tilde{\mathbf{H}}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\alpha}})/\tilde{\sigma}_0^2$$

i.e.,

$$\operatorname{tr}(\tilde{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i') = (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\alpha}})'\tilde{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\tilde{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\alpha}}) \ ,$$

and on using (3.22) this is

$$\operatorname{tr}(\tilde{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i') = \mathbf{y}'\tilde{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\tilde{\mathbf{P}}\mathbf{y}, \quad \text{for } i = 1, \cdots, c. \tag{4.18}$$

Define

$$\boldsymbol{\Delta}_i = \operatorname{diag}\{\mathbf{0} \cdots \mathbf{0} \ \ \mathbf{I}_{q_i} \ \ \mathbf{0} \cdots \mathbf{0}\} = \boldsymbol{\Delta}_i^2 \ , \tag{4.19}$$

a null matrix except for $\mathbf{I}_{q_i}$ on the diagonal in the i'th row of submatrices. Also, recall from (3.6)

$$\mathbf{T}^{*} = (\mathbf{I} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D})^{-1} \ ,$$

of order $\sum_{i=1}^{c} q_i$. On partitioning $\mathbf{T}^{*}$ into c^2 matrices, this means

$$\begin{bmatrix} \mathbf{T}^{*}_{11} & \cdots & \mathbf{T}^{*}_{1c} \\ \vdots & & \\ \mathbf{T}^{*}_{c1} & \cdots & \mathbf{T}^{*}_{cc} \end{bmatrix} \begin{bmatrix} \mathbf{I} + \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{Z}_1\sigma_1^2 & \cdots & \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{Z}_c\sigma_c^2 \\ \vdots & & \vdots \\ \mathbf{Z}_c'\mathbf{R}^{-1}\mathbf{Z}_1\sigma_1^2 & \cdots & \mathbf{I} + \mathbf{Z}_c'\mathbf{R}^{-1}\mathbf{Z}_c\sigma_c^2 \end{bmatrix} = \mathbf{I} \tag{4.20}$$

so that

$$\mathbf{T}^{*}_{ii} + \sigma_i^2 \sum_{j=1}^{c} \mathbf{T}^{*}_{ij}\mathbf{Z}_j'\mathbf{R}^{-1}\mathbf{Z}_i = \mathbf{I}_{q_i}$$

and (4.21)

$$\mathbf{T}^{*}_{ik} + \sigma_k^2 \sum_{j=1}^{c} \mathbf{T}^{*}_{ij}\mathbf{Z}_j'\mathbf{R}^{-1}\mathbf{Z}_k = \mathbf{0} \quad \text{for } i \neq k.$$

Then a rewriting of the left-hand side of (4.18) can be based on the following development:

$$\begin{aligned}
\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') &= \mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i) \\
&= \mathrm{tr}(\underset{\sim}{\Delta}_i\underset{\sim}{Z}'\underset{\sim}{V}^{-1}\underset{\sim}{Z}\underset{\sim}{\Delta}_i), \quad \text{using (4.19)} \\
&= \mathrm{tr}(\underset{\sim}{\Delta}_i\underset{\sim}{T}^*\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{\Delta}_i), \text{ using (3.11)} \qquad (4.22) \\
&= \mathrm{tr}(\text{i'th diagonal submatrix of } \underset{\sim}{T}^*\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}) \\
&= \mathrm{tr}(\sum_{j=1}^{c}\underset{\sim}{T}^*_{ij}\underset{\sim}{Z}_j'\underset{\sim}{R}^{-1}\underset{\sim}{Z}_i) \\
&= \mathrm{tr}(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}^*_{ii})/\sigma_i^2 \quad \text{from (4.21).} \qquad (4.23)
\end{aligned}$$

Also, the right-hand side of (4.18) is based on

$$\begin{aligned}
\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y} &= (\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y})'\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y} \\
&= (\underset{\sim}{\Delta}_i\underset{\sim}{Z}'\underset{\sim}{P}\underset{\sim}{y})'\underset{\sim}{\Delta}_i\underset{\sim}{Z}'\underset{\sim}{P}\underset{\sim}{y} \\
&= (\underset{\sim}{\Delta}_i\tilde{\underset{\sim}{v}})'\underset{\sim}{\Delta}_i\tilde{\underset{\sim}{v}}, \quad \text{from (3.25)} \\
&= \tilde{\underset{\sim}{v}}_i'\tilde{\underset{\sim}{v}}_i \\
&= \tilde{\underset{\sim}{b}}_i'\tilde{\underset{\sim}{b}}_i/\sigma_i^4 . \qquad (4.24)
\end{aligned}$$

Substituting (4.23) and (4.24) into the ML equations (4.18) reduces them to

$$[q_i - \mathrm{tr}(\underset{\sim}{T}^*_{ii})]\tilde{\sigma}_i^2 = \tilde{\underset{\sim}{b}}_i'\tilde{\underset{\sim}{b}}_i$$

with the two equivalent forms

$$\tilde{\sigma}_i^2 = [\tilde{\underset{\sim}{b}}_i'\tilde{\underset{\sim}{b}}_i + \tilde{\sigma}_i^2\mathrm{tr}(\underset{\sim}{T}^*_{ii})]/q_i \qquad (4.24a)$$

and

$$\tilde{\sigma}_i^2 = \tilde{\underset{\sim}{b}}_i'\tilde{\underset{\sim}{b}}_i/[q_i - \mathrm{tr}(\underset{\sim}{T}^*_{ii})] . \qquad (4.24b)$$

Together with (4.17) for $\tilde{\sigma}_0^2$ these provide iterative procedures in the following manner, where the superscript (r) indicates values calculated after the r'th round of iteration for $\tilde{\underset{\sim}{\sigma}}^2$, with r = 0 denoting the starting values.

$$\tilde{\sigma}_0^{2(r+1)} = \frac{\underset{\sim}{y}'[\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\alpha}}^{(r)} - \underset{\sim}{Z}\tilde{\underset{\sim}{b}}^{(r)}]}{N} \tag{4.25}$$

and, for i = 1, ⋯, c

$$\tilde{\sigma}_i^{2(r+1)} = \frac{(\tilde{\underset{\sim}{b}}_i'\tilde{\underset{\sim}{b}}_i)^{(r)} + \tilde{\sigma}_i^{2(r)}\mathrm{tr}(\underset{\sim}{T}_{ii}^{*(r)})}{q_i} \tag{4.26}$$

or

$$\tilde{\sigma}_i^{2(r+1)} = \frac{(\tilde{\underset{\sim}{b}}_i'\tilde{\underset{\sim}{b}}_i)^{(r)}}{q_i - \mathrm{tr}(\underset{\sim}{T}_{ii}^{*(r)})} \tag{4.27}$$

Equations (4.25), (4.26) and (4.27) are [H6.2], [H6.1] and [H6.5], respectively.

Comment: Expression (4.26) originated in Henderson [1973]. It and (4.27) always give positive estimates because their denominators are positive, except in trivial cases. Evidence of this comes from a lemma of Harville's [1975], part of which is as follows.

Harville's Lemma: (i) $\mathrm{tr}(\underset{\sim}{T}_{ii}^{*})$ is positive.

(ii) $q_i \geq \mathrm{tr}(\underset{\sim}{T}_{ii}^{*})$ for $\sigma_i^2 > 0$ with strict inequality holding if $\underset{\sim}{Z}_i \neq \underset{\sim}{0}$.

Proof: From (3.6), $\underset{\sim}{T}^{*-1} = \underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}$. Observe that $\underset{\sim}{D}^{\frac{1}{2}}$ exists, for $\underset{\sim}{D} = \underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{D}^{\frac{1}{2}}$ and define

$$\underset{\sim}{W} = \underset{\sim}{I} + \underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{D}^{\frac{1}{2}} . \tag{4.28}$$

Then, by a simple extension of Searle [1971b, p. 24, Lemma 8], $\underset{\sim}{W}$ is positive definite. Furthermore, by inspection,

$$\underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{T}^{*-1} = \underset{\sim}{W}\underset{\sim}{D}^{\frac{1}{2}}$$

and so

$$\underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{T}^{*} = \underset{\sim}{W}^{-1}\underset{\sim}{D}^{\frac{1}{2}} \tag{4.29}$$

with $\underset{\sim}{W}^{-1}$ existing because of positive definiteness; and for the same reason its diagonal elements are positive. With $\underset{\sim}{D}$ and hence $\underset{\sim}{D}^{\frac{1}{2}}$ being diagonal, (4.29) therefore indicates that $\underset{\sim}{T}^{*}$ has positive diagonal elements. Hence $\mathrm{tr}(\underset{\sim}{T}_{ii}) > 0$, so proving (i).

From (4.23)

$$\begin{aligned}[q_i - \mathrm{tr}(\underset{\sim}{T}^{*}_{ii})]/\sigma_i^2 &= \mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') \\ &= \mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i) \\ &= \mathrm{var}(\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{y}) , \\ &\geq 0 .\end{aligned}$$

Therefore, providing $\sigma_i^2 > 0$, $q_i \geq \mathrm{tr}(\underset{\sim}{T}^{*}_{ii})$ with strict inequality holding if $\underset{\sim}{Z}_i \neq \underset{\sim}{0}$. <u>Q.E.D.</u>

The importance of this lemma is that (4.25) with either (4.26) or (4.27) always gives positive values for the estimates. This valuable feature of this iterative procedure is discussed at length in Harville [1977]. Note, however, that although (4.26) and (4.27) each give only positive values, they will not necessarily yield the same sequence of iterates.

4.4. THE INFORMATION MATRIX

In Chapter 3 we provide an extensive description of alternative solutions to alternative ways of writing the MME's. The compendium-like nature of that description is provided with a view to having some of the many alternatives explicitly available, if for no other reason, to save readers the effort of pursuing these alternatives themselves. We adopt the same attitude here for information matrices. They are developed for $\dot{\boldsymbol{\sigma}}^2$ and $\dot{\boldsymbol{\gamma}}$ [in equations (4.33) and (4.37), respectively], together with computing expressions for them in terms of the MME's [equations (4.60) and (4.61)]. The likelihood equations are also written in terms of these information matrices, equations (4.70) and (4.73). Expressions are also given for use in the Newton-Raphson algorithm [(4.85) through (4.89), and (4.94)], and use of the information matrices in the Fisher scoring algorithm is also indicated.

In terms of estimating $\boldsymbol{\theta}$, as in section 4.2, the information matrix for $\boldsymbol{\theta}$, to be denoted by $\mathbf{I}(\boldsymbol{\theta})$, is

$$\mathbf{I}(\boldsymbol{\theta}) = -E\left\{\frac{\partial^2 L}{\partial\theta_i \partial\theta_j}\right\} = E\,\frac{\partial L}{\partial\boldsymbol{\theta}}\left(\frac{\partial L}{\partial\boldsymbol{\theta}}\right)' , \tag{4.30}$$

the second equality coming from Lemma 2. The importance of $\mathbf{I}(\boldsymbol{\theta})$ is that, in terms of the n implicit in L of Lemmas 1 and 2,

$$\lim_{n\to\infty} \operatorname{var}(\tilde{\boldsymbol{\theta}}) = -\frac{1}{n}[\mathbf{I}(\boldsymbol{\theta})]^{-1} \tag{4.31}$$

where $\tilde{\boldsymbol{\theta}}$ is the ML estimator of $\boldsymbol{\theta}$. In variance components models there is some difficulty in exactly defining what is meant by a limit as $n\to\infty$ for (4.31), a difficulty which has been considered by Hartley and Rao [1967] and, more recently, by Miller [1973].

a. For $\dot{\underset{\sim}{\sigma}}^2$

Using the likelihood function (4.1), Searle [1970] has shown that the information matrix for $\dot{\underset{\sim}{\sigma}}^2$ of (1.27) is

$$\underset{\sim}{I}\begin{bmatrix}\underset{\sim}{\alpha}\\ \dot{\underset{\sim}{\sigma}}^2\end{bmatrix} = \tfrac{1}{2}\begin{bmatrix}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X} & \underset{\sim}{0}\\ \underset{\sim}{0} & \underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)\end{bmatrix}$$

where

$$\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2) = \tfrac{1}{2}\left\{\mathrm{tr}\left(\underset{\sim}{V}^{-1}\frac{\partial\underset{\sim}{V}}{\partial\sigma_i^2}\underset{\sim}{V}^{-1}\frac{\partial\underset{\sim}{V}}{\partial\sigma_j^2}\right)\right\} \tag{4.32}$$

$$= \tfrac{1}{2}\left\{\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\right\}, \qquad \text{for } i,j = 0, 1, \cdots, c. \tag{4.33}$$

The large sample variances are obtained from this by inversion

$$\mathrm{var}(\tilde{\underset{\sim}{\sigma}}^2) = [I(\dot{\sigma}^2)]^{-1} .$$

b. For $\dot{\underset{\sim}{\gamma}}$

To obtain the information matrix for σ_0^2 and $\underset{\sim}{\gamma}$, we use the middle expression of (4.30), differentiate (4.9) and (4.10) and (after changing sign) take expectations: From (4.9)

$$\frac{\partial^2 L}{\partial(\gamma_0)^2} = \frac{N}{2\sigma_0^4} - \frac{(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})'\underset{\sim}{H}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})}{\sigma_0^6}$$

and then

$$-E\,\frac{\partial^2 L}{\partial(\gamma_0)^2} = \frac{-N}{2\sigma_0^4} + \frac{\mathrm{tr}[\underset{\sim}{H}^{-1}E(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})']}{\sigma_0^6}$$

$$= \frac{-N}{2\sigma_0^4} + \frac{\mathrm{tr}(\underset{\sim}{H}^{-1}\sigma_0^2\underset{\sim}{H})}{\sigma_0^6} = \frac{N}{2\sigma_0^4} . \tag{4.34}$$

Similarly from (4.10)

$$\frac{\partial^2 L}{\partial \gamma_i \partial \gamma_0} = \frac{-(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})}{2\sigma_0^4}$$

so that

$$-E\,\frac{\partial^2 L}{\partial \gamma_i \partial \gamma_0} = \frac{\mathrm{tr}[\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}E(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})']}{2\sigma_0^4}$$

$$= \frac{\mathrm{tr}(\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}\sigma_0^2\underset{\sim}{H})}{2\sigma_0^4} = \frac{\mathrm{tr}\ \underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i}{2\sigma_0^2}\ . \qquad (4.35)$$

And also from (4.10)

$$\frac{\partial^2 L}{\partial \gamma_i \partial \gamma_j} = (-1/2\sigma_0^2)[-\sigma_0^2\ \mathrm{tr}(\underset{\sim}{H}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i')$$

$$+ (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})'(\underset{\sim}{H}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1} + \underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{H}^{-1})(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha})]$$

with

$$-E\,\frac{\partial^2 L}{\partial \gamma_i \partial \gamma_j} = -\tfrac{1}{2}\mathrm{tr}(\underset{\sim}{H}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') + \frac{2}{2\sigma_0^2}\ \mathrm{tr}(\underset{\sim}{H}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}\underset{\sim}{H}\sigma_0^2)$$

$$= \tfrac{1}{2}\mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i)\ . \qquad (4.36)$$

Repeated use of (2.2), (2.3), (2.4), (2.11) and (2.12) is made in deriving the results (4.34), (4.35) and (4.36). Assembled into a matrix these results give $\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}})$ for $\dot{\underset{\sim}{\gamma}}' = [\gamma_0 \quad \underset{\sim}{\gamma}']$ of (1.27) as

$$\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}}) = \underset{\sim}{I}\begin{pmatrix}\gamma_0 \\ \underset{\sim}{\gamma}\end{pmatrix} = \frac{1}{2}\begin{bmatrix} N/\sigma_0^4 & \{\mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i)\}'/\sigma_0^2 \\ \text{sym.} & \{\mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i)\}\end{bmatrix} \qquad (4.37)$$

for $i, j = 1, 2, \cdots, c$.

$\{\mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}\underset{\sim}{Z}_i)\}'$ in (4.37) is a row vector of c elements; its column form also occurs in (4.37), where the abbreviation "sym." is shown, signifying that the matrix is symmetric. This abbreviation is used repeatedly in what follows.

Then the large sample sampling variances are

$$\mathrm{var}(\tilde{\dot{\underset{\sim}{\gamma}}}) = [\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}})]^{-1} .$$

c. Transformed parameters

The information matrices of (4.33) and (4.37) are related in the manner indicated by the following lemma.

Lemma 4.3. When parameters $\underset{\sim}{\theta}$ are transformed to $\underset{\sim}{\Delta}$ by a one-to-one transformation, the information matrix $\underset{\sim}{I}(\underset{\sim}{\Delta})$ is given in terms of $\underset{\sim}{I}(\underset{\sim}{\theta})$ by

$$\underset{\sim}{I}(\underset{\sim}{\Delta}) = \underset{\sim}{J}'\underset{\sim}{I}(\underset{\sim}{\theta})\underset{\sim}{J} \tag{4.38}$$

where $\underset{\sim}{J}$ is defined as

$$\underset{\sim}{J} = \underset{\sim}{J}_{\underset{\sim}{\theta}:\underset{\sim}{\Delta}} = \{m_{ij}\} = \left\{\frac{\partial\theta_i}{\partial\Delta_j}\right\} \quad \text{for } i,j = 1, \cdots, k, \tag{4.39}$$

the Jacobian matrix of $\underset{\sim}{\theta}$ with respect to $\underset{\sim}{\Delta}$ (Zacks [1971, 5.1.15, p. 227]).

Proof: The logarithm of the likelihood, L, can be thought of as a function of either $\underset{\sim}{\theta}$ or $\underset{\sim}{\Delta}$, whichever parameter set is being used. Whatever form is used it represents the same thing. Then, from (4.30)

$$\underset{\sim}{I}(\underset{\sim}{\Delta}) = E\,\frac{\partial L}{\partial\underset{\sim}{\Delta}}\left(\frac{\partial L}{\partial\underset{\sim}{\Delta}}\right)'$$

and, in view of the functional relationship between $\underset{\sim}{\Delta}$ and $\underset{\sim}{\theta}$,

$$\frac{\partial L}{\partial\Delta} = \left\{\frac{\partial L}{\partial\Delta_i}\right\} \quad \text{for } i = 1, \cdots, k,$$

$$= \left\{\sum_{j=1}^{k}\frac{\partial L}{\partial\theta_j}\frac{\partial\theta_j}{\partial\Delta_i}\right\} = \left\{\sum_{j=1}^{k}\frac{\partial\theta_j}{\partial\Delta_i}\frac{\partial L}{\partial\theta_j}\right\} \quad \text{for } i = 1, \cdots, k, \quad = \underset{\sim}{J}'\,\frac{\partial L}{\partial\underset{\sim}{\theta}} .$$

Hence

$$\underset{\sim}{I}(\underset{\sim}{\Delta}) = E\ \underset{\sim}{J}'\ \frac{\partial L}{\partial \underset{\sim}{\theta}}\left(\underset{\sim}{J}'\ \frac{\partial L}{\partial \underset{\sim}{\theta}}\right)' = \underset{\sim}{J}'E\left[\frac{\partial L}{\partial \underset{\sim}{\theta}}\left(\frac{\partial L}{\partial \underset{\sim}{\theta}}\right)'\right]\underset{\sim}{J} = \underset{\sim}{J}'\underset{\sim}{I}(\underset{\sim}{\theta})\underset{\sim}{J}. \qquad \underline{\text{Q.E.D.}}$$

We illustrate (4.38) by using it to obtain $\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}})$ of (4.37) from $\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)$ of (4.32). First, from (1.27)

$$\dot{\underset{\sim}{\gamma}} = \begin{bmatrix} \sigma_0^2 \\ \underset{\sim}{\gamma} \end{bmatrix} = \begin{bmatrix} 1 & \underset{\sim}{0} \\ \underset{\sim}{0} & (1/\sigma_0^2)\underset{\sim}{I} \end{bmatrix}\dot{\underset{\sim}{\sigma}}^2 , \tag{4.40}$$

so that with $\sigma_0^2 \equiv \gamma_0$

$$\dot{\underset{\sim}{\sigma}}^2 = \begin{bmatrix} 1 & \underset{\sim}{0} \\ \underset{\sim}{0} & \sigma_0^2\underset{\sim}{I} \end{bmatrix}\dot{\underset{\sim}{\gamma}} = \begin{bmatrix} \gamma_0 \\ \gamma_0\gamma_1 \\ \vdots \\ \gamma_0\gamma_c \end{bmatrix} .$$

Hence from (4.39)

$$\underset{\sim}{J} = \underset{\sim}{J}_{\dot{\underset{\sim}{\sigma}}^2:\dot{\underset{\sim}{\gamma}}} = \left\{\frac{\partial \sigma_i^2}{\partial \gamma_j}\right\} \quad \text{for } i,j = 0,\ 1,\ 2,\ \cdots,\ c, \quad = \begin{bmatrix} 1 & \underset{\sim}{0} \\ \underset{\sim}{\gamma} & \sigma_0^2\underset{\sim}{I} \end{bmatrix} . \tag{4.41}$$

Then, on substituting (4.33) and (4.41) into (4.38), and using $\underset{\sim}{Z}_0 = \underset{\sim}{I}$ of (1.23) in doing so, we get

$$\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}}) = \frac{1}{2}\begin{bmatrix} 1 & \underset{\sim}{\gamma}' \\ \underset{\sim}{0} & \sigma_0^2\underset{\sim}{I} \end{bmatrix}\begin{bmatrix} \operatorname{tr}(\underset{\sim}{V}^{-2}) & \{\operatorname{tr}(\underset{\sim}{V}^{-2}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\}' \\ \text{sym.} & \{\operatorname{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\} \end{bmatrix}\begin{bmatrix} 1 & \underset{\sim}{0} \\ \underset{\sim}{\gamma} & \sigma_0^2\underset{\sim}{I} \end{bmatrix}$$

for $i,j = 1,\ 2,\ \cdots,\ c,$

$$\equiv \frac{1}{2}\begin{bmatrix} f_0 & \underset{\sim}{f}' \\ \underset{\sim}{f} & \underset{\sim}{F} \end{bmatrix}, \text{ say} \tag{4.42}$$

where

$$f_0 = \mathrm{tr}(\underset{\sim}{V}^{-2}) + 2\mathrm{tr}(\underset{\sim}{V}^{-2}\sum_{i=1}^{c}\gamma_i \underset{\sim}{Z}_i \underset{\sim}{Z}_i') + \sum_{i=1}^{c}\sum_{j=1}^{c}\mathrm{tr}(\underset{\sim}{V}^{-1}\gamma_i \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \underset{\sim}{V}^{-1}\gamma_j \underset{\sim}{Z}_j \underset{\sim}{Z}_j')$$

$$= \mathrm{tr}(\underset{\sim}{V}^{-2}) + 2\mathrm{tr}[\underset{\sim}{V}^{-2}(\underset{\sim}{V} - \sigma_0^2 \underset{\sim}{I})/\sigma_0^2] + \mathrm{tr}[\underset{\sim}{V}^{-1}(\underset{\sim}{V} - \sigma_0^2\underset{\sim}{I})/\sigma_0^2]^2, \text{ using (1.30)}$$

$$= \mathrm{tr}(\underset{\sim}{V}^{-2} + 2\underset{\sim}{V}^{-1}/\sigma_0^2 - 2\underset{\sim}{V}^{-2} + \underset{\sim}{I}/\sigma_0^4 - 2\underset{\sim}{V}^{-1}/\sigma_0^2 + \underset{\sim}{V}^{-2})$$

$$= N/\sigma_0^4 \tag{4.43}$$

and, with $\underset{\sim}{f}' = \{f_j\}$ for $j = 1, \cdots, c$,

$$f_j = \sigma_0^2[\mathrm{tr}(\underset{\sim}{V}^{-2}\underset{\sim}{Z}_j\underset{\sim}{Z}_j') + \sum_i \gamma_i \mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')]$$

$$= \sigma_0^2 \mathrm{tr}(\underset{\sim}{V}^{-2}\underset{\sim}{Z}_j\underset{\sim}{Z}_j') + \mathrm{tr}[\underset{\sim}{V}^{-1}(\underset{\sim}{V} - \sigma_0^2\underset{\sim}{I})\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j']$$

$$= \mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\ ; \tag{4.44}$$

and

$$\underset{\sim}{F} = \sigma_0^4\{\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\} \quad \text{for } i,j = 1, 2, \cdots, c. \tag{4.45}$$

Using (4.43), (4.44) and (4.45) in (4.42) together with $\underset{\sim}{V}^{-1} = \underset{\sim}{H}^{-1}/\sigma_0^2$ gives (4.37) exactly.

d. Relationships between sampling variances

Computing methodology for maximum likelihood estimation is usually based on deriving estimators of $\dot{\underset{\sim}{\gamma}}$ rather than of $\dot{\underset{\sim}{\sigma}}^2$ directly. (See Hartley and Rao [1967], Hemmerle and Hartley [1973] and Corbeil and Searle [1976a].) The same computing effort can also yield estimates of $\mathrm{var}(\tilde{\sigma}_0^2)$ and $\mathrm{var}(\tilde{\underset{\sim}{\gamma}})$. However, although $\tilde{\underset{\sim}{\sigma}}^2$ is easily obtainable from $\tilde{\underset{\sim}{\gamma}}$, the derivation of $\mathrm{var}(\tilde{\underset{\sim}{\sigma}}^2)$ from $\mathrm{var}(\tilde{\underset{\sim}{\gamma}})$ is a little more involved. It comes from (4.38), using $\underset{\sim}{J}$ of (4.41):

$$\text{var}(\tilde{\sigma}^2) = [\mathbf{I}(\dot{\sigma}^2)]^{-1} = [\mathbf{J}'^{-1}\mathbf{I}(\dot{\gamma})\mathbf{J}^{-1}]^{-1}$$

$$= \mathbf{J}\text{var}(\tilde{\gamma})\mathbf{J}'$$

$$= \begin{bmatrix} 1 & \mathbf{0} \\ \boldsymbol{\gamma} & \sigma^2\mathbf{I} \end{bmatrix} \begin{bmatrix} v(\tilde{\sigma}_0^2) & \text{cov}(\tilde{\sigma}_0^2, \tilde{\boldsymbol{\gamma}}') \\ \text{sym.} & \text{var}(\tilde{\boldsymbol{\gamma}}) \end{bmatrix} \begin{bmatrix} 1 & \boldsymbol{\gamma}' \\ \mathbf{0} & \sigma_0^2\mathbf{I} \end{bmatrix}$$

$$= \begin{bmatrix} v(\tilde{\sigma}_0^2) & v(\tilde{\sigma}_0^2)\boldsymbol{\gamma}' + \sigma_0^2 \,\text{cov}(\tilde{\sigma}_0^2, \tilde{\boldsymbol{\gamma}}') \\ \text{sym.} & \mathbf{G} \end{bmatrix}$$

with

$$\mathbf{G} = v(\tilde{\sigma}_0^2)\boldsymbol{\gamma}\boldsymbol{\gamma}' + \sigma_0^2[\text{cov}(\tilde{\sigma}_0^2, \boldsymbol{\gamma})\boldsymbol{\gamma}' + \boldsymbol{\gamma}\,\text{cov}(\tilde{\sigma}_0^2, \tilde{\boldsymbol{\gamma}}')] + \sigma_0^4 \,\text{var}(\tilde{\boldsymbol{\gamma}}).$$

Hence, apart from $v(\tilde{\sigma}_0^2)$, typical terms are

$$v(\tilde{\sigma}_i^2) = \gamma_i^2 \, v(\tilde{\sigma}_0^2) + 2\sigma_0^2\gamma_i \,\text{cov}(\tilde{\sigma}_0^2, \tilde{\gamma}_i) + \sigma_0^4 \, v(\tilde{\gamma}_i) \,, \tag{4.46}$$

$$\text{cov}(\tilde{\sigma}_0^2, \tilde{\sigma}_i^2) = \gamma_i \, v(\tilde{\sigma}_0^2) + \sigma_0^2 \,\text{cov}(\tilde{\sigma}^2, \tilde{\gamma}_i) \tag{4.47}$$

and

$$\text{cov}(\tilde{\sigma}_i^2, \tilde{\sigma}_j^2) = \gamma_i\gamma_j \, v(\tilde{\sigma}_0^2) + \sigma_0^2[\gamma_j \,\text{cov}(\tilde{\sigma}_0^2, \tilde{\gamma}_i) + \gamma_i \,\text{cov}(\tilde{\sigma}_0^2, \tilde{\gamma}_j)] + \sigma_0^4 \,\text{cov}(\tilde{\gamma}_i, \tilde{\gamma}_j). \tag{4.48}$$

These are the results given in (45) of Corbeil and Searle [1976a].

e. Relationships with MME's

We show here the relationships between information matrices and the MME's.

(i) For $\dot{\sigma}^2$ From (4.33)

$$\mathbf{I}(\dot{\sigma}^2) = \frac{1}{2}\begin{bmatrix} \text{tr}(\mathbf{V}^{-2}) & \{\text{tr}(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1})\}' \\ \text{sym.} & \{\text{tr}(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j')\} \end{bmatrix} \quad \text{for } i,j = 1, \cdots, c. \tag{4.49}$$

Based upon (4.21), (4.23) and methodology described for REML by Harville [1977, Sec. 5] we show how to express the matrices in (4.49) in terms of $\underset{\sim}{T}^*$. First, for the upper right-hand term,

$$\begin{aligned}
\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}) &= \mathrm{tr}(\underset{\sim}{\Delta}_i\underset{\sim}{Z}'\underset{\sim}{V}^{-1}\underset{\sim}{V}^{-1}\underset{\sim}{Z}\underset{\sim}{\Delta}_i), \text{ using (4.19)} \\
&= \mathrm{tr}(\underset{\sim}{\Delta}_i\underset{\sim}{T}^*\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{T}^{*\prime}\underset{\sim}{\Delta}_i), \text{ from (3.11)} \\
&= \mathrm{tr}(\text{i'th diagonal submatrix of } \underset{\sim}{T}^*\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}\underset{\sim}{T}^{*\prime})/\sigma_0^2, \text{ using } \underset{\sim}{R} = \sigma_0^2\underset{\sim}{I} \\
&= \mathrm{tr}\Big(\sum_{k=1}^{c}\sum_{j=1}^{c}\underset{\sim}{T}^*_{ij}\underset{\sim}{Z}_j'\underset{\sim}{R}^{-1}\underset{\sim}{Z}_k\underset{\sim}{T}^{*\prime}_{ik}\Big)/\sigma_0^2 \\
&= \mathrm{tr}\Big(\sum_{j=1}^{c}\underset{\sim}{T}^*_{ij}\underset{\sim}{Z}_j'\underset{\sim}{R}^{-1}\underset{\sim}{Z}_i\underset{\sim}{T}^{*\prime}_{ii} + \sum_{k\neq i}^{c}\sum_{j}\underset{\sim}{T}^*_{ij}\underset{\sim}{Z}_j'\underset{\sim}{R}^{-1}\underset{\sim}{Z}_k\underset{\sim}{T}^{*\prime}_{ik}\Big)/\sigma_0^2 \\
&= \mathrm{tr}\Big[(\underset{\sim}{I} - \underset{\sim}{T}^*_{ii})\underset{\sim}{T}^{*\prime}_{ii}/\sigma_i^2 + \sum_{k\neq i}^{c}(-\underset{\sim}{T}^*_{ik}/\sigma_k^2)\underset{\sim}{T}^{*\prime}_{ik}\Big]/\sigma_0^2 \qquad \text{from (4.21).}
\end{aligned}$$

But from (3.7)

$$\sigma_k^2 T^*_{ki} = (T^{*\prime})_{ki}\sigma_i^2 = (T^*_{ik})'\sigma_i^2 \equiv T^{*\prime}_{ik}\sigma_i^2 \,.$$

Therefore

$$\underset{\sim}{T}^{*\prime}_{ik}/\sigma_k^2 = \underset{\sim}{T}^*_{ki}/\sigma_i^2$$

and so

$$\begin{aligned}
\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}) &= \mathrm{tr}\Big[(\underset{\sim}{I} - \underset{\sim}{T}^*_{ii})\underset{\sim}{T}^*_{ii} - \sum_{k\neq i}^{c}\underset{\sim}{T}^*_{ik}\underset{\sim}{T}^*_{ki}\Big]/\sigma_i^2\sigma_0^2 \\
&= \Big[\mathrm{tr}(\underset{\sim}{T}^*_{ii}) - \sum_{k=1}^{c}\mathrm{tr}(\underset{\sim}{T}^*_{ik}\underset{\sim}{T}^*_{ki})\Big]/\sigma_i^2\sigma_0^2 \,. \qquad (4.50)
\end{aligned}$$

Then, for the lower right-hand term of (4.49) there are two kinds of submatrices:

$$\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') = \mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i)^2 = \mathrm{tr}(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}^*_{ii})^2/\sigma_i^4 \,, \qquad (4.51)$$

from (4.23); and for $i \neq j$

$$\mathrm{tr}(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j') = \mathrm{tr}(\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_i)$$

$$= \mathrm{tr}(\mathbf{\Delta}_i\mathbf{Z}'\mathbf{V}^{-1}\mathbf{Z}\mathbf{\Delta}_j\mathbf{\Delta}_j\mathbf{Z}'\mathbf{V}^{-1}\mathbf{Z}\mathbf{\Delta}_i) \tag{4.52}$$

where

$$\mathbf{\Delta}_i\mathbf{Z}'\mathbf{V}^{-1}\mathbf{Z}\mathbf{\Delta}_j = i,j\text{'th submatrix of } \mathbf{Z}'\mathbf{V}^{-1}\mathbf{Z}$$

$$= i,j\text{'th submatrix of } \mathbf{T}^*\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z},\text{ from (3.11)}$$

$$= \sum_{k=1}^{c} \mathbf{T}^*_{ik}\mathbf{Z}_k'\mathbf{R}^{-1}\mathbf{Z}_j$$

$$= -\mathbf{T}^*_{ij}/\sigma_j^2\text{ , from (4.21).} \tag{4.53}$$

Hence (4.52) is, for $i \neq j$

$$\mathrm{tr}(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j') = \mathrm{tr}(\mathbf{T}^*_{ij}\mathbf{T}^{*\prime}_{ij})/\sigma_j^4 = \mathrm{tr}(\mathbf{T}^*_{ij}\mathbf{T}^*_{ji})/\sigma_i^2\sigma_j^2\ , \tag{4.54}$$

this last step coming from the result that precedes (4.50). Finally we consider $\mathrm{tr}(\mathbf{V}^{-2})$ and, to do so, rely on assuming

$$\mathbf{R} = \sigma_0^2\mathbf{I} \tag{4.55}$$

as is usually the case in variance components models. Then from (2.14) and (3.6)

$$\mathbf{V}^{-2} = (\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}\mathbf{D}\mathbf{T}^*\mathbf{Z}'\mathbf{R}^{-1})^2 \tag{4.56}$$

$$= (\mathbf{I} - \mathbf{Z}\mathbf{D}\mathbf{T}^*\mathbf{Z}'\mathbf{R}^{-1})^2/\sigma_0^4\ , \tag{4.57}$$

from (4.55), so that

$$\sigma_0^4\mathbf{V}^{-2} = \mathbf{I} - 2\mathbf{Z}\mathbf{D}\mathbf{T}^*\mathbf{Z}'\mathbf{R}^{-1} + \mathbf{Z}\mathbf{D}\mathbf{T}^*\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D}\mathbf{T}^*\mathbf{Z}'\mathbf{R}^{-1}$$

$$= \mathbf{I} - 2\mathbf{Z}\mathbf{D}\mathbf{T}^*\mathbf{Z}'\mathbf{R}^{-1} + \mathbf{Z}\mathbf{D}\mathbf{T}^*(\mathbf{T}^{*-1} - \mathbf{I})\mathbf{T}^*\mathbf{Z}'\mathbf{R}^{-1},\text{ from (3.6),}$$

$$= \mathbf{I} - \mathbf{Z}\mathbf{D}\mathbf{T}^*\mathbf{Z}'\mathbf{R}^{-1} - \mathbf{Z}\mathbf{D}\mathbf{T}^{*2}\mathbf{Z}'\mathbf{R}^{-1}\ . \tag{4.58}$$

Hence

$$\sigma_0^4 \mathrm{tr}(\underset{\sim}{V}^{-2}) = N - \mathrm{tr}[\underset{\sim}{T}^{*}(\underset{\sim}{T}^{*-1} - \underset{\sim}{I}) + \underset{\sim}{T}^{*2}(\underset{\sim}{T}^{*-1} - \underset{\sim}{I})]$$

$$= N - \mathrm{tr}(\underset{\sim}{I} - \underset{\sim}{T}^{*2}) .$$

Therefore

$$\mathrm{tr}(\underset{\sim}{V}^{-2}) = \frac{N - q}{\sigma_0^4} + \frac{1}{\sigma_0^4} \sum_i \sum_j \mathrm{tr}(\underset{\sim}{T}^{*}_{ij} \underset{\sim}{T}^{*}_{ji}) . \qquad (4.59)$$

Substituting (4.50), (4.51), (4.54) and (4.59) into (4.49) gives

$$\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2) = \frac{1}{2} \begin{bmatrix} \dfrac{N-q}{\sigma_0^4} + \dfrac{1}{\sigma_0^4} \sum\limits_i^c \sum\limits_j^c \mathrm{tr}(\underset{\sim}{T}^{*}_{ij} \underset{\sim}{T}^{*}_{ji}) & \left\{[\mathrm{tr}(\underset{\sim}{T}^{*}_{ii}) - \sum\limits_{k=1}^{c} \mathrm{tr}(\underset{\sim}{T}^{*}_{ik} \underset{\sim}{T}^{*}_{ki})]/\sigma_0^2 \sigma_i^2\right\} \\ \text{sym.} & \left\{\begin{matrix} \text{diagonal submatrices: } \mathrm{tr}(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}^{*}_{ii})^2/\sigma_i^4 \\ \text{off-diag. submatrices: } \mathrm{tr}(\underset{\sim}{T}^{*}_{ij} \underset{\sim}{T}^{*}_{ji})/\sigma_i^2 \sigma_j^2 \end{matrix}\right\} \end{bmatrix} .$$

(4.60)

(ii) For $\dot{\gamma}$ From (4.37)

$$\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}}) = \frac{1}{2} \begin{bmatrix} N/\sigma_0^4 & \{\mathrm{tr}(\underset{\sim}{Z}_i' \underset{\sim}{H}^{-1} \underset{\sim}{Z}_i)\}'/\sigma_0^2 \\ \text{sym.} & \{\mathrm{tr}(\underset{\sim}{Z}_i' \underset{\sim}{H}^{-1} \underset{\sim}{Z}_j \underset{\sim}{Z}_j' \underset{\sim}{H}^{-1} \underset{\sim}{Z}_i)\} \end{bmatrix} . \qquad (4.37)$$

For the upper right-hand term

$$\mathrm{tr}(\underset{\sim}{Z}_i' \underset{\sim}{H}^{-1} \underset{\sim}{Z}_i)/\sigma_0^2 = \mathrm{tr}(\underset{\sim}{\Delta}_i \underset{\sim}{Z}' \underset{\sim}{V}^{-1} \underset{\sim}{Z} \underset{\sim}{\Delta}_i) = [q_i - \mathrm{tr}(\underset{\sim}{T}^{*}_{ii})]/\sigma_i^2$$

from (4.23); and the lower right-hand term is the same as that of (4.60) multiplied by σ_0^4 . Hence

$$\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}}) = \frac{1}{2}\begin{bmatrix} N/\sigma_0^4 & \left\{[q_i - \mathrm{tr}(\underset{\sim}{T}^{*}_{ii})]/\sigma_i^2\right\}' \\ \text{sym.} & \begin{cases} \text{diagonal submatrices: } \mathrm{tr}(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}^{*}_{ii})^2/\gamma_i^2 \\ \text{off-diag. submatrices: } \mathrm{tr}(\underset{\sim}{T}^{*}_{ij}\underset{\sim}{T}^{*}_{ji})/\gamma_i\gamma_j \end{cases} \end{bmatrix} . \tag{4.61}$$

f. Equations for estimators

The ML equations for $\dot{\underset{\sim}{\sigma}}^2$ and $\dot{\underset{\sim}{\gamma}}$ can be written in a manner that utilizes the form of the information matrices.

(i) For $\dot{\underset{\sim}{\sigma}}^2$ The equations for estimating $\dot{\underset{\sim}{\sigma}}^2$ are (4.6), the left-hand side of which is

$$\begin{aligned} \mathrm{tr}(\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') &\equiv \mathrm{tr}(\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{V}}^{-1}\tilde{\underset{\sim}{V}}) \\ &= \mathrm{tr}(\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{V}}^{-1}\sum_{j=0}^{c}\tilde{\sigma}_j^2\underset{\sim}{Z}_j\underset{\sim}{Z}_j'), \text{ using (1.25)} \\ &= \sum_{j=0}^{c}[\mathrm{tr}(\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')]\tilde{\sigma}_j^2 . \end{aligned} \tag{4.62}$$

Hence (4.6) is

$$\sum_{j=0}^{c}[\mathrm{tr}(\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j']\tilde{\sigma}_j^2 = \underset{\sim}{y}'\tilde{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{P}}\underset{\sim}{y}\,, \quad \text{for } i = 0, \cdots, c.$$

These can be written in vector form as

$$\{\mathrm{tr}(\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\}\tilde{\underset{\sim}{\sigma}}^2 = \{\underset{\sim}{y}'\tilde{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{P}}\underset{\sim}{y}\}, \quad \text{for } i = 0, \cdots, c. \tag{4.63}$$

Comparing the left-hand side with (4.33) we see that the matrix multiplying $\tilde{\underset{\sim}{\sigma}}$ is $2\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)$ with $\underset{\sim}{V}^{-1}$ replaced by $\tilde{\underset{\sim}{V}}^{-1}$. On making this replacement we write $\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)$ as $\underset{\sim}{I}(\tilde{\underset{\sim}{\sigma}}^2)$ and (4.63) becomes

$$\begin{aligned} [\underset{\sim}{I}(\tilde{\underset{\sim}{\sigma}}^2)]\tilde{\underset{\sim}{\sigma}}^2 &= \tfrac{1}{2}\{\underset{\sim}{y}'\tilde{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{P}}\underset{\sim}{y}\} \\ &= \tfrac{1}{2}\{(\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\alpha}})'\tilde{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\tilde{\underset{\sim}{V}}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\alpha}})\} \quad \text{for } i = 0, 1, \cdots, c. \end{aligned} \tag{4.64}$$

This is equivalent to [H8.1].

Equations (4.64) can be expressed in terms of the MME's. The coefficient matrix, $[\mathbf{I}(\tilde{\boldsymbol{\sigma}}^2)]$, on the left-hand side is available from (4.60), using $\tilde{\boldsymbol{\sigma}}^2$ in place of $\dot{\boldsymbol{\sigma}}^2$. Hence we have only to express the right-hand side of (4.64) in terms of the MME's. To do this we make a simplification in the notation.

Notation: For notational convenience we drop the ~'s from (4.64), save in the vector $\tilde{\boldsymbol{\sigma}}^2$ to be solved for, and consider the right-hand sides as

$$\tfrac{1}{2}[\mathbf{y}'\mathbf{P}^2\mathbf{y} \quad \{\mathbf{y}'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{y}\}]' \quad \text{for } i = 1, \cdots, c. \tag{4.65}$$

We now develop the two terms in (4.65), again adopting the assumption $\mathbf{R} = \sigma_0^2\mathbf{I}$ in (4.55). From (2.48) and (3.8)

$$\mathbf{P}^2 = (\mathbf{S} - \mathbf{SZDTZ'S})^2 \tag{4.66}$$

and on using $\mathbf{R} = \sigma_0^2\mathbf{I}$ so that, from (2.52)

$$\mathbf{S} = \mathbf{M}/\sigma_0^2 \quad \text{and} \quad \mathbf{S}^2 = \mathbf{M}^2/\sigma_0^4 = \mathbf{M}/\sigma_0^4 = \mathbf{S}/\sigma_0^2\ ,$$

we get

$$\begin{aligned}
\mathbf{P}^2 &= \mathbf{S}^2 - \mathbf{S}^2\mathbf{ZDTZ'S} - \mathbf{SZDTZ'S}^2 + \mathbf{SZDTZ'S}^2\mathbf{ZDTZ'S} \\
&= [\mathbf{S} - 2\mathbf{SZDTZ'S} + \mathbf{SZDT}(\mathbf{T}^{-1} - \mathbf{I})\mathbf{TZ'S}]/\sigma_0^2, \text{ using (3.8)} \\
&= (\mathbf{S} - \mathbf{SZDTZ'S} - \mathbf{SZDT}^2\mathbf{Z'S})/\sigma_0^2\ .
\end{aligned} \tag{4.67}$$

Then

$$\begin{aligned}
\mathbf{y}'\mathbf{P}^2\mathbf{y} &= (\mathbf{y}'\mathbf{Sy} - \mathbf{y}'\mathbf{SZDTZ'Sy} - \mathbf{y}'\mathbf{SZDT}^2\mathbf{Z}'\mathbf{Sy})/\sigma_0^2 \\
&= (\mathbf{y}'\mathbf{Sy} - \mathbf{y}'\mathbf{SZD}\tilde{\boldsymbol{\nu}} - \mathbf{y}'\mathbf{SZDT}\tilde{\boldsymbol{\nu}})/\sigma_0^2, \text{ from (3.14) and (3.26)} \\
&= (\mathbf{y}'\mathbf{Sy} - \mathbf{y}'\mathbf{SZ}\tilde{\mathbf{b}} - \mathbf{y}'\mathbf{SZT'D}\tilde{\boldsymbol{\nu}})/\sigma_0^2, \text{ from (3.5) and (3.9)} \\
&= [\mathbf{y}'\mathbf{S}(\mathbf{y} - \mathbf{Z}\tilde{\mathbf{b}}) - \tilde{\boldsymbol{\nu}}'\mathbf{D}\tilde{\boldsymbol{\nu}}]/\sigma_0^2, \text{ from (3.14) and (3.25)}
\end{aligned}$$

$$= [\underset{\sim}{y}'\underset{\sim}{R}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\alpha}} - \underset{\sim}{Z}\tilde{\underset{\sim}{b}}) - \sum_{i=1}^{c} \tilde{\underset{\sim}{v}}_i'\tilde{\underset{\sim}{v}}_i\sigma_i^2]/\sigma_0^2, \text{ using (3.30)}$$

$$= \underset{\sim}{y}'(\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\alpha}} - \underset{\sim}{Z}\tilde{\underset{\sim}{b}})/\sigma_0^4 - \frac{1}{\sigma_0^2}\sum_{i=1}^{c}\frac{\tilde{\underset{\sim}{b}}_i'\tilde{\underset{\sim}{b}}_i}{\sigma_i^2} . \tag{4.68}$$

Finally, for $i = 1, \cdots, c$, we have from (4.24)

$$\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y} = \tilde{\underset{\sim}{v}}_i'\tilde{\underset{\sim}{v}}_i = \tilde{\underset{\sim}{b}}_i'\tilde{\underset{\sim}{b}}_i/\sigma_i^4 . \tag{4.69}$$

Equations (4.60) with $\tilde{\underset{\sim}{\sigma}}^2$ in place of $\underset{\sim}{\sigma}^2$, on the left-hand side of (4.64) and in (4.68) and (4.69) for the right-hand side, therefore provide an iterative procedure for estimating $\dot{\underset{\sim}{\sigma}}^2$:

$$\begin{bmatrix} \dfrac{N-q}{\tilde{\sigma}_0^4} + \dfrac{1}{\tilde{\sigma}_0^4}\sum\limits_{i=1}^{c}\sum\limits_{j=1}^{c}\operatorname{tr}(\tilde{\underset{\sim}{T}}^{*}_{ij}\tilde{\underset{\sim}{T}}^{*}_{ji}) & \left\{[\operatorname{tr}(\tilde{\underset{\sim}{T}}^{*}_{ii}) - \sum\limits_{k=1}^{c}\operatorname{tr}(\tilde{\underset{\sim}{T}}^{*}_{ik}\tilde{\underset{\sim}{T}}^{*}_{ki})]/\tilde{\sigma}_0^2\tilde{\sigma}_i^2\right\} \\ \text{sym.} & \left\{\begin{matrix}\text{diagonal submatrices: } \operatorname{tr}(\underset{\sim}{I}_{q_i} - \tilde{\underset{\sim}{T}}^{*}_{ii})^2/\tilde{\sigma}_i^4 \\ \text{off-diag. submatrices: } \operatorname{tr}(\tilde{\underset{\sim}{T}}^{*}_{ij}\tilde{\underset{\sim}{T}}^{*}_{ji})/\tilde{\sigma}_i^2\tilde{\sigma}_j^2\end{matrix}\right\} \end{bmatrix} \begin{bmatrix} \tilde{\sigma}_0^2 \\ \\ \tilde{\underset{\sim}{\sigma}}^2 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} \underset{\sim}{y}'(\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\alpha}} - \underset{\sim}{Z}\tilde{\tilde{\underset{\sim}{b}}})/\sigma_0^4 - \dfrac{1}{\tilde{\sigma}_0^2}\sum\limits_{i=1}^{c}\tilde{\tilde{\underset{\sim}{b}}}_i'\tilde{\tilde{\underset{\sim}{b}}}_i/\tilde{\sigma}_i^2 \\ \\ \{\tilde{\tilde{b}}_i'\tilde{\tilde{b}}_i/\tilde{\sigma}_i^4\} \end{bmatrix} \tag{4.70}$$

for $i, j = 1, \cdots, c$.

The symbol $\tilde{\tilde{\underset{\sim}{b}}}_i$ indicates $\tilde{\underset{\sim}{b}}_i$ from the MME's, calculated using $\tilde{\underset{\sim}{V}}$ in place of $\underset{\sim}{V}$; in the same way that $\tilde{\underset{\sim}{\alpha}}$ is $\hat{\underset{\sim}{\alpha}}$ with $\tilde{\underset{\sim}{V}}$ in place of $\underset{\sim}{V}$.

Clearly, as an iterative procedure this is an alternative to but not as simple as (4.25) with either (4.26) or (4.27). Expressions similar to those on the left-hand side of (4.70), i.e., akin to those in $\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)$ of (4.60), were developed by Henderson [1973] and have been commented on by Schaeffer [1976].

(ii) For $\underset{\sim}{\gamma}$ Equation (4.13) is the basis for estimating $\underset{\sim}{\gamma}$; its left-hand side is

$$\begin{aligned}
\mathrm{tr}(\tilde{H}^{-1}Z_iZ_i') &\equiv \mathrm{tr}(\tilde{H}^{-1}Z_iZ_i'\tilde{H}^{-1}\tilde{H}) \\
&= \mathrm{tr}[\tilde{H}^{-1}Z_iZ_i'\tilde{H}^{-1}(I + \sum_{i=1}^{c} Z_jZ_j'\gamma_j)] \\
&= \mathrm{tr}(\tilde{H}^{-1}Z_iZ_i'\tilde{H}^{-1}) + \sum_{i=1}^{c} \mathrm{tr}(\tilde{H}^{-1}Z_iZ_i'\tilde{H}^{-1}Z_jZ_j')\gamma_j \qquad (4.71) \\
&= \mathrm{tr}(\tilde{H}^{-1}Z_iZ_i'\tilde{H}^{-1}) + [I(\tilde{\gamma})]\tilde{\gamma} ,
\end{aligned}$$

on comparing the last term in (4.71) with the lower right-hand submatrix of (4.37). Hence the ML equations (4.13) for $\underset{\sim}{\gamma}$ are

$$[I(\tilde{\gamma})]\tilde{\gamma} = \left\{(y - X\tilde{\alpha})'\tilde{H}^{-1}Z_iZ_i'\tilde{H}^{-1}(y - X\tilde{\alpha})/\tilde{\sigma}_0^2 - \mathrm{tr}(\tilde{H}^{-1}Z_iZ_i'\tilde{H}^{-1})\right\} \qquad (4.72)$$

for $i = 1, \cdots, c,$

$$= \tilde{\sigma}_0^2 y'\tilde{P}Z_iZ_i'\tilde{P}y - \tilde{\sigma}_0^4\ \mathrm{tr}(\tilde{V}^{-1}Z_iZ_i'\tilde{V}^{-1}) .$$

From (4.61), (4.24) and (4.50) this can be written as

$$\left\{\begin{array}{l}\text{diagonal submatrices: } \mathrm{tr}(I_{q_i} - \tilde{T}^*_{ii})^2/\tilde{\gamma}_i^2 \\ \text{off-diag. submatrices: } \mathrm{tr}(\tilde{T}^*_{ij}\tilde{T}^*_{ji})/\tilde{\gamma}_i\tilde{\gamma}_j\end{array}\right\}\tilde{\gamma}$$

$$= \frac{1}{2}\left\{\frac{1}{\sigma_0^2}\,\frac{\tilde{b}_i'\tilde{b}_i}{\gamma_i^2} - [\mathrm{tr}(T^*_{ii}) - \sum_{k=1}^{c}\mathrm{tr}(T^*_{ik}T^*_{ki})]/\tilde{\gamma}_i\right\} \qquad (4.73)$$

for $i,j = 1, \cdots, c.$

Since (4.25) with either (4.26) or (4.27) are probably more appealing than these derivations, we pursue them no further.

4.5. COMPUTING ALGORITHMS

It is definitely not the purpose of this Notebook to discuss computing difficulties involved in obtaining solutions to equations such as (4.5), (4.6) or (4.11), (4.12) and (4.13), or (4.70). Clearly, no analytic solutions to these equations are possible and so one resorts to iterative, arithmetic methods. Two well-known procedures for this are the Newton-Raphson algorithm and the Fisher scoring algorithm, for each of which we here give a bare outline.

a. Newton-Raphson

Suppose θ is the vector of parameters to be estimated and L is the logarithm of the likelihood. Then if θ_0 is the value of θ at the end of any round of iteration and θ_1 is the value at the end of the next round, the Newton-Raphson algorithm is to calculate θ_1 as

$$\theta_1 = \theta_0 - \left(\frac{\partial^2 L}{\partial\theta\partial\theta'}\right)^{-1} \frac{\partial L}{\partial\theta}\Bigg|_{\theta=\theta_0} . \tag{4.74}$$

(i) For σ^2 The elements of the matrix to be inverted here, for $\theta = \sigma^2$, come from differentiating

$$\frac{\partial L}{\partial\sigma_i^2} = -\tfrac{1}{2}\mathrm{tr}(V^{-1}Z_iZ_i') + \tfrac{1}{2}(y - X\alpha)'V^{-1}Z_iZ_i'V^{-1}(y - X\alpha) \tag{4.75}$$

obtained from (4.3) after using $\partial V/\partial\sigma_i^2 = Z_iZ_i'$. Thus

$$\begin{aligned}\frac{\partial^2 L}{\partial\sigma_i^2\partial\sigma_j^2} &= \tfrac{1}{2}\mathrm{tr}(V^{-1}Z_jZ_j'V^{-1}Z_iZ_i') \\ &\quad + \tfrac{1}{2}(y - X\alpha)'[-V^{-1}Z_jZ_j'V^{-1}Z_iZ_i'V^{-1} - V^{-1}Z_iZ_i'V^{-1}Z_jZ_j'V^{-1}](y - X\alpha) \\ &= \tfrac{1}{2}\mathrm{tr}(V^{-1}Z_iZ_i'V^{-1}Z_jZ_j') - (y - X\alpha)'V^{-1}Z_iZ_i'V^{-1}Z_jZ_j'V^{-1}(y - X\alpha) .\end{aligned}$$

Since, in any round of iteration, this expression will be used in (4.74) with parameters replaced by their computed estimates obtained from the preceding round, one effect of this is to replace $\underset{\sim}{\alpha}$ by $\hat{\underset{\sim}{\alpha}}$. Making this replacement gives

$$\left.\frac{\partial^2 L}{\partial\sigma_i^2 \partial\sigma_j^2}\right|_{\alpha=\hat{\alpha}} = \tfrac{1}{2}\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j') - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{P}\underset{\sim}{y} \qquad (4.76)$$

for $i, j = 0, \cdots, c$.

We consider four forms of this:

$$\left.\frac{\partial^2 L}{\partial(\sigma_0^2)^2}\right|_{\alpha=\hat{\alpha}} = \tfrac{1}{2}\mathrm{tr}(\underset{\sim}{V}^{-2}) - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{V}^{-1}\underset{\sim}{P}\underset{\sim}{y}\,, \qquad (4.77)$$

$$\left.\frac{\partial^2 L}{\partial\sigma_0^2 \partial\sigma_i^2}\right|_{\alpha=\hat{\alpha}} = \tfrac{1}{2}\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}) - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}\,, \qquad (4.78)$$

for $i = 1, \cdots, c$,

$$\left.\frac{\partial^2 L}{\partial(\sigma_i^2)^2}\right|_{\alpha=\hat{\alpha}} = \tfrac{1}{2}\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}\,, \qquad (4.79)$$

for $i = 1, \cdots, c$,

and

$$\left.\frac{\partial^2 L}{\partial\sigma_i^2 \partial\sigma_j^2}\right|_{\alpha=\hat{\alpha}} = \tfrac{1}{2}\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j') - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{P}\underset{\sim}{y}\,, \qquad (4.80)$$

for $i \neq j = 1, \cdots, c$.

Computing formulae for the first terms in each of these expressions, in terms of $\underset{\sim}{T}^*_{ij}$ and $\underset{\sim}{\tilde{v}}_i'$ or $\underset{\sim}{\tilde{b}}_i'$, are available in (4.59), (4.50), (4.51) and (4.54), respectively. The second terms for (4.77) - (4.80) are as follows. From (3.10) with $\underset{\sim}{R} = \sigma_0^2\underset{\sim}{I}$,

$$\begin{aligned}
\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{V}^{-1}\underset{\sim}{P}\underset{\sim}{y} &= \underset{\sim}{y}'\underset{\sim}{P}^2\underset{\sim}{y}/\sigma_0^2 - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{T}^*\underset{\sim}{Z}'\underset{\sim}{P}\underset{\sim}{y}/\sigma_0^4 \\
&= \underset{\sim}{y}'\underset{\sim}{P}^2\underset{\sim}{y}/\sigma_0^2 - \underset{\sim}{\tilde{v}}'\underset{\sim}{D}\underset{\sim}{T}^*\underset{\sim}{\tilde{v}}/\sigma_0^4 \\
&= \frac{\underset{\sim}{y}'(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\alpha} - \underset{\sim}{Z}\underset{\sim}{b})}{\sigma_0^6} - \frac{\sum_{i=1}^{c}\sigma_i^2\underset{\sim}{\tilde{v}}_i'\underset{\sim}{\tilde{v}}_i}{\sigma_0^4} - \frac{\sum_{i=1}^{c}\sigma_i^2(\underset{\sim}{\tilde{v}}_i'\sum_{j=1}^{c}\underset{\sim}{T}^*_{ij}\underset{\sim}{\tilde{v}}_j)}{\sigma_0^4}\,.
\end{aligned} \qquad (4.81)$$

For (4.78)

$$\begin{aligned}
y'PV^{-1}Z_iZ_i'Py &= y'PZ_iZ_i'V^{-1}Py \\
&= y'PZ\Delta_i\Delta_i Z'V^{-1}Py \\
&= \nu'\Delta_i\Delta_i T^* Z'R^{-1}Py, \text{ from (3.25) and (3.11)} \\
&= \nu_i'\Delta_i T^* Z'Py/\sigma_0^2 , \\
&= \nu_i'\Delta_i T^* \nu/\sigma_0^2 , \text{ from (3.25)} \\
&= \tilde{\nu}_i'(\sum_{j=1}^{c} T^*_{ij}\tilde{\nu}_j)/\sigma_0^2 .
\end{aligned} \tag{4.82}$$

For (4.79)

$$\begin{aligned}
y'PZ_iZ_i'V^{-1}Z_iZ_i'Py &= y'PZ\Delta_i(\Delta_i Z'V^{-1}Z\Delta_i)\Delta_i Z'Py \\
&= \nu'\Delta_i[(I - T^*_{ii})/\sigma_i^2]\Delta_i\nu , \text{ from (3.25) and (4.23)} \\
&= (\nu_i'\nu_i - \nu_i'T^*_{ii}\tilde{\nu}_i)/\sigma_i^2 .
\end{aligned} \tag{4.83}$$

And finally, for (4.80), with $i \neq j$

$$\begin{aligned}
y'PZ_iZ_i'V^{-1}Z_jZ_j'Py &= y'PZ\Delta_i(\Delta_i Z'V^{-1}Z\Delta_j)\Delta_j Z'Py \\
&= \nu'\Delta_i(-T^*_{ij}/\sigma_j^2)\Delta_j\tilde{\nu} , \text{ from (3.25) and (4.53)} \\
&= \tilde{\nu}_i'T^*_{ij}\tilde{\nu}_j/\sigma_j^2 .
\end{aligned} \tag{4.84}$$

Hence on assembling (4.67), (4.62) - (4.64) and (4.81) - (4.84) into (4.77) - (4.81) we get the matrix to be inverted in the Newton-Raphson procedure, (4.74), as

$$\left.\frac{\partial^2 L}{\partial\dot{\sigma}^2(\partial\dot{\sigma}^2)'}\right|_{\alpha=\hat{\alpha}} = \begin{bmatrix} e_0 & e' \\ e & E \end{bmatrix} \tag{4.85}$$

where, from putting (4.67) and (4.81) into (4.77)

$$e_0 = \frac{(N-q)}{2\sigma_0^4} + \frac{1}{2\sigma_0^4}\sum_{i=1}^{c}\sum_{j=1}^{c}\mathrm{tr}(T^*_{ij}T^*_{ji}) - \frac{\underset{\sim}{y}'(\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\alpha}} - \underset{\sim}{Z}\tilde{\underset{\sim}{b}})}{\sigma_0^6} + \frac{\sum_{i=1}^{c}\sigma_i^2\tilde{\underset{\sim}{\nu}}_i'(\tilde{\underset{\sim}{\nu}}_i + \sum_{j=1}^{c}\underset{\sim}{T}^*_{ij}\tilde{\underset{\sim}{\nu}}_j)}{\sigma_0^4}\;; \qquad (4.86)$$

from putting (4.62) and (4.82) into (4.78)

$$\underset{\sim}{e} = \{e_i\} \text{ with } e_i = \frac{\mathrm{tr}(\underset{\sim}{T}^*_{ii})}{2\sigma_0^2\sigma_i^2} - \frac{\sum_{k=1}^{c}\mathrm{tr}(\underset{\sim}{T}^*_{ik}\underset{\sim}{T}^*_{ki})}{2\sigma_0^2\sigma_i^2} - \frac{\underset{\sim}{\nu}_i'(\sum_{j=1}^{c}\underset{\sim}{T}^*_{ij}\underset{\sim}{\nu}_j)}{\sigma_0^2} \qquad (4.87)$$

for $i = 1, \cdots, c,$

and with (4.63) and (4.83) going into (4.79), and (4.64) and (4.84) into (4.80)

$\underset{\sim}{E} = \{e_{ij}\}$

$$\text{with } e_{ii} = \frac{\mathrm{tr}(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}^*_{ii})^2}{2\sigma_i^2} - \frac{\tilde{\underset{\sim}{\nu}}_i'(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}^*_{ii})\underset{\sim}{\nu}_i}{\sigma_i^2} \qquad (4.88)$$

$$\text{and } e_{ij} = \frac{\mathrm{tr}(\underset{\sim}{T}^*_{ij}\underset{\sim}{T}^*_{ji})}{2\sigma_i^2\sigma_j^2} + \frac{\tilde{\underset{\sim}{\nu}}_i'\underset{\sim}{T}^*_{ij}\tilde{\underset{\sim}{\nu}}_j}{\sigma_j^2}\,, \quad \text{for } i \neq j = 1, \cdots, c. \qquad (4.89)$$

(ii) For $\hat{\underset{\sim}{\gamma}}$ For using Newton-Raphson on $\underset{\sim}{\gamma}$, we find in the derivation of (4.34) - (4.36) that, corresponding to (4.77) through (4.80),

$$\left.\frac{\partial^2 L}{\partial\gamma_0^2}\right|_{\underset{\sim}{\alpha}=\hat{\underset{\sim}{\alpha}}} = \frac{\frac{1}{2}N}{\sigma_0^4} - \frac{\underset{\sim}{y}'\underset{\sim}{PVH}^{-1}\underset{\sim}{PVy}}{\sigma_0^6} = (\tfrac{1}{2}N - \underset{\sim}{y}'\underset{\sim}{Py})/\sigma_0^4 \qquad (4.90)$$

$$\left.\frac{\partial^2 L}{\partial\gamma_0\partial\gamma_i}\right|_{\underset{\sim}{\alpha}=\hat{\underset{\sim}{\alpha}}} = -\underset{\sim}{y}'\underset{\sim}{PVH}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{H}^{-1}\underset{\sim}{VPy}/2\sigma_0^4$$

$$= -\underset{\sim}{y}'\underset{\sim}{PZ}_i\underset{\sim}{Z}_i'\underset{\sim}{Py} \qquad (4.91)$$

$$\left.\frac{\partial^2 L}{\partial\gamma_i\partial\gamma_j}\right|_{\underset{\sim}{\alpha}=\hat{\underset{\sim}{\alpha}}} = \tfrac{1}{2}\sigma_0^4[\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j') - 2(\underset{\sim}{y} - \underset{\sim}{X\alpha})'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim}{X\alpha})]$$

so that

$$\left.\frac{\partial^2 L}{\partial \gamma_i^2}\right|_{\underset{\sim}{\alpha}=\underset{\sim}{\hat{\alpha}}} = \tfrac{1}{2}\sigma_0^4[\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i')^2 - 2\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}] \tag{4.92}$$

and

$$\left.\frac{\partial^2 L}{\partial \gamma_i \partial \gamma_j}\right|_{\underset{\sim}{\alpha}=\underset{\sim}{\hat{\alpha}}} = \tfrac{1}{2}\sigma_0^4[\mathrm{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j') - 2\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{P}\underset{\sim}{y}] \;. \tag{4.93}$$

Then on using (3.33) in (4.90), (4.69) in (4.91), (4.51) and (4.83) in (4.92), and (4.54) and (4.84) in (4.93) we get

$$\left.\frac{\partial^2 L}{\partial \gamma_i \partial \gamma_j}\right|_{\underset{\sim}{\alpha}=\underset{\sim}{\hat{\alpha}}}$$

$$= \begin{bmatrix} \tfrac{1}{2}\gamma_0^{-2}[N - 2\underset{\sim}{y}'\underset{\sim}{S}/\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\tilde{b}})] & -\tfrac{1}{2}\{\underset{\sim}{\tilde{\nu}}_i'\underset{\sim}{\tilde{\nu}}_i\}' \\ \text{sym.} & \left\{\begin{array}{l} \text{diagonal terms: } \tfrac{1}{2}\gamma_i^{-2}\mathrm{tr}(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}^{*}_{ii})^2 - \gamma_0\gamma_i^{-1}\underset{\sim}{\tilde{\nu}}_i'(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}^{*}_{ii})\underset{\sim}{\tilde{\nu}}_i \\ \text{off-diag. terms: } \tfrac{1}{2}\gamma_i^{-1}\gamma_j^{-1}\mathrm{tr}(\underset{\sim}{T}^{*}_{ij}\underset{\sim}{T}^{*}_{ji}) + \gamma_0\gamma_j^{-1}\underset{\sim}{\tilde{\nu}}_i'\underset{\sim}{T}^{*}_{ij}\underset{\sim}{\tilde{\nu}}_j \end{array}\right\} \end{bmatrix}$$

$$\text{for } i,j = 1, \cdots, c. \tag{4.94}$$

The matrix in the lower right-hand segment of (4.94) is $\sigma_0^4\underset{\sim}{E}$ for $\underset{\sim}{E}$ defined in (4.88) and (4.89), and used in (4.85).

b. The Fisher scoring method

This method derives $\underset{\sim}{\theta}_1$ as

$$\underset{\sim}{\theta}_1 = \underset{\sim}{\theta}_0 + \left(-E\,\frac{\partial^2 L}{\partial\underset{\sim}{\theta}\partial\underset{\sim}{\theta}'}\right)^{-1} \left.\frac{\partial L}{\partial \underset{\sim}{\theta}}\right|_{\underset{\sim}{\theta}=\underset{\sim}{\theta}_0} \;. \tag{4.95}$$

The matrix to be inverted here is the inverse of the information matrix, as can be seen from (4.30). Various expressions for this are available in (4.33), (4.49), (4.60) and (4.37), and (4.61). In view of Lemma 2 of Sec. 4.3, this can also be written as

$$\underset{\sim}{\theta}_1 = \underset{\sim}{\theta}_0 + \left[E\left(\frac{\partial L}{\partial \underset{\sim}{\theta}}\right)\left(\frac{\partial L}{\partial \underset{\sim}{\theta}}\right)'\right]^{-1} \frac{\partial L}{\partial \underset{\sim}{\theta}}\Bigg|_{\underset{\sim}{\theta}=\underset{\sim}{\theta}_0} . \tag{4.96}$$

In all cases, in the variance components situation, the non-negativity constraints discussed earlier must also be made part of these and all computing procedures. Excellent discussion of numerical properties of these and of their relationship to Hemmerle and Hartley's [1973] algorithm is given in Jennrich and Sampson [1976]. Harville [1975, 1977, Sec. 6] also discusses these algorithms as well as some suggested by Anderson [1970], Henderson [1973] and others.

Chapter 5

REML: RESTRICTED MAXIMUM LIKELIHOOD

Thompson [1962], for balanced data and the completely random model, suggested estimating variance components by maximizing that portion of the likelihood which is invariant to the mean. Patterson and Thompson [1971] extended this to the mixed model, and for the randomized block design with unequal block sizes proposed a method that can be extended to any unbalanced data situation. Corbeil and Searle [1976a] give specific algorithms for carrying this out, and Harville [1975, 1977] discusses general properties.

The main results derived in this chapter are indicated in the following table.

Table 5.1. Equation Numbers in Chapter 5 of the Main Results for Restricted Maximum Likelihood (REML)

Parameter	Estimation Equations		
	Basic equations	Equations using MME's	Equations for $\hat{\hat{\sigma}}^2$
σ^2	(5.17) - (5.18)	(5.36) - (5.38)	(5.19)
γ	(5.21) - (5.24)		

Parameter	Information Matrices			
	Basic equations	Using MME's	Used in estimation equations with MME's	Used in Newton-Raphson equations
σ^2	(5.39), (5.40)	(5.51)	(5.56)	(5.57)
γ	(5.50)	(5.52)		

The term "error contrast" is used by Harville [1974, 1977] to denote any linear function $\underset{\sim}{k}'\underset{\sim}{y}$ of the observations that that has zero expectation, i.e., $E(\underset{\sim}{k}'\underset{\sim}{y}) = \underset{\sim}{k}'\underset{\sim}{X}\underset{\sim}{\alpha} = \underset{\sim}{0}$ and hence $\underset{\sim}{k}'\underset{\sim}{X} = \underset{\sim}{0}$. More generally we confine attention to a set of such contrasts, $\underset{\sim}{K}'\underset{\sim}{y}$ with $\underset{\sim}{K}'\underset{\sim}{X} = \underset{\sim}{0}$ and from the discussion following (2.7) observe that with $\underset{\sim}{X}$ having N rows and rank p^*, there are N columns in $\underset{\sim}{K}'$ and it has rank not greater than $N - p^*$. There is obviously no merit in dealing with $\underset{\sim}{K}'\underset{\sim}{y}$ if some rows of $\underset{\sim}{K}'$ (and hence elements of $\underset{\sim}{K}'\underset{\sim}{y}$) are linear combinations of others; neither should we lose information by using a $\underset{\sim}{K}'\underset{\sim}{y}$ that has fewer elements than the possible maximum. We therefore deal with $(\underset{\sim}{K}')_{(N-p^*)\times N}$ of full row rank $N - p^*$ and $\underset{\sim}{K}'\underset{\sim}{X} = \underset{\sim}{0}$ as in (2.72).

We first show that it does not matter what matrix $\underset{\sim}{K}'$ of this specification we use; the differentiable part of the log likelihood is the same for all $\underset{\sim}{K}'$'s. Further, the log likelihood can be written without even involving $\underset{\sim}{K}'$ explicitly.

5.1. INVARIANCE OF THE RESTRICTED LIKELIHOOD

Estimation by restricted maximum likelihood entails estimating $\underset{\sim}{\sigma}^2$ by maximizing the likelihood of $\underset{\sim}{K}'\underset{\sim}{y}$ for $\underset{\sim}{K}'$ as just specified. The logarithm of this likelihood is, similar to (4.1),

$$L(\underset{\sim}{K}'\underset{\sim}{y}) = -\tfrac{1}{2}(N - p^*)\log 2\pi - \tfrac{1}{2}\log|\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K}| - \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K}'\underset{\sim}{y} .$$

But, by Lemma 2.4, $\underset{\sim}{K}' = \underset{\sim}{F}'\underset{\sim}{A}'$ for $\underset{\sim}{F}$ non-singular and $\underset{\sim}{A}$ defined in (2.22), and so from Lemma 2.8 and also using (2.75)

$$L(\underset{\sim}{K}'\underset{\sim}{y}) = -\tfrac{1}{2}(N - p^*)\log 2\pi - \log|\underset{\sim}{F}| - \tfrac{1}{2}\log|\underset{\sim}{A}'\underset{\sim}{V}\underset{\sim}{A}| - \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{y} .$$

From (2.76) this can also be expressed as

$$\begin{aligned} L(\underset{\sim}{K}'\underset{\sim}{y}) = &-\tfrac{1}{2}(N - p^*)\log 2\pi - \log|\underset{\sim}{F}| + \tfrac{1}{2}\log|\underset{\sim}{X}^{*\prime}\underset{\sim}{X}^*| \\ &- \tfrac{1}{2}\log|\underset{\sim}{V}| - \tfrac{1}{2}\log|\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^*| - \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{y} . \end{aligned} \tag{5.1}$$

The first three terms in (5.1) depend on $\mathbf{X}$, $\mathbf{A}$ and $\mathbf{K}$ but not on any parameters of the model. Hence, apart from an additive constant the log likelihood of $\mathbf{K}'\mathbf{y}$ for <u>any</u> $\mathbf{K}'$ specified by (2.72) is

$$L_1 = -\tfrac{1}{2}\log|\mathbf{V}| - \tfrac{1}{2}\log|\mathbf{X}^{*\prime}\mathbf{V}^{-1}\mathbf{X}^{*}| - \tfrac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y} \tag{5.2}$$

$$= -\tfrac{1}{2}\log|\mathbf{V}| - \tfrac{1}{2}\log|\mathbf{X}^{*\prime}\mathbf{V}^{-1}\mathbf{X}^{*}| - \tfrac{1}{2}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}})\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}}) , \tag{5.3}$$

using (3.32). This expression is given in Harville [1977, p. 325].

Conclusions of practical importance to be drawn from (5.2) are

(i) for every set of $N - p^*$ LIN error contrasts the log likelihood is the same,

and

(ii) apart from a constant that is inconsequential to the estimation of $\boldsymbol{\sigma}^2$, that log likelihood is L_1 of (5.1). This means that no matter what $N - p^*$ linearly independent error contrasts are used, maximizing their likelihood always lead to the same equations for estimating $\boldsymbol{\sigma}^2$.

5.2. EQUIVALENT EXPRESSIONS FOR THE LOG LIKELIHOOD

We here develop a variety of expressions for L_1, some of which are given following [H5.3]. They utilize

$$\mathbf{C}^* = \begin{bmatrix} \mathbf{X}^{*\prime}\mathbf{R}^{-1}\mathbf{X}^* & \mathbf{X}^{*\prime}\mathbf{R}^{-1}\mathbf{Z}\mathbf{D} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}^* & \mathbf{I} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D} \end{bmatrix} \tag{5.3}$$

which is $\mathbf{C}$ of (2.62) with $\mathbf{X}^*$ replacing $\mathbf{X}$. We begin with expressions for the determinant of $\mathbf{C}^*$. The first equality in (2.5) gives

$$|\underset{\sim}{C}^*| = |\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^*|\ |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{ZD} - \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{X}^*(\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^*)^{-1}\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{ZD}|$$

$$= |\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^*|\ |\underset{\sim}{I} + \underset{\sim}{Z}'\{\underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{X}^*(\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^*)^{-1}\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\}\underset{\sim}{ZD}|$$

$$= |\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^*|\ |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{SZD}|\ , \quad \text{from (2.36) and (2.59);} \tag{5.4}$$

and the second equality in (2.5) gives rise to

$$|\underset{\sim}{C}^*| = |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{ZD}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^* - \underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{ZD}(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{ZD})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{X}^*|$$

$$= |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{ZD}|\ |\underset{\sim}{X}^{*\prime}\{\underset{\sim}{R}^{-1} - \underset{\sim}{R}^{-1}\underset{\sim}{ZD}(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{ZD})^{-1}\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\}\underset{\sim}{X}^*|$$

$$= |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{ZD}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^*|\ , \quad \text{from (2.14).} \tag{5.5}$$

Therefore

$$|\underset{\sim}{C}^*| = |\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^*|\ |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{SZD}| = |\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^*|\ |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{ZD}|\ . \tag{5.6}$$

Also,

$$|\underset{\sim}{V}| = |\underset{\sim}{ZDZ}' + \underset{\sim}{R}| = |\underset{\sim}{R}|\ |\underset{\sim}{I} + \underset{\sim}{R}^{-1}\underset{\sim}{ZDZ}'| = |\underset{\sim}{R}|\ |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{ZD}|\ , \tag{5.7}$$

on using (2.7), and so (5.6) gives

$$|\underset{\sim}{R}|\ |\underset{\sim}{C}^*| = |\underset{\sim}{V}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^*| = |\underset{\sim}{R}|\ |\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^*|\ |\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{SZD}|\ . \tag{5.8}$$

Using (5.8) in (5.2), together with (3.32) and (3.34) for $\underset{\sim}{y}'\underset{\sim}{Py}$ gives L_1 of (5.2) as

$$L_1 = -\tfrac{1}{2}\log|\underset{\sim}{R}| - \tfrac{1}{2}\log|\underset{\sim}{C}^*| - \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{R}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\tilde{\alpha} - \underset{\sim}{Z}\tilde{b})\ , \tag{5.9}$$

which is the first equation in the right-hand column of page 326 of Harville [1977]. Similarly, the second equation there comes from applying (5.6) and (3.33) to (5.9) to yield

$$L_1 = -\tfrac{1}{2}\log|\underset{\sim}{R}| - \tfrac{1}{2}\log|\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^{*}| - \tfrac{1}{2}\log|\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{SZD}| - \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{S}(\underset{\sim}{y} - \underset{\sim}{Z}\tilde{\underset{\sim}{b}}) \qquad (5.10)$$

or, on using (3.32) again, it is

$$L_1 = -\tfrac{1}{2}\log|\underset{\sim}{R}| - \tfrac{1}{2}\log|\underset{\sim}{X}^{*\prime}\underset{\sim}{R}^{-1}\underset{\sim}{X}^{*}| - \tfrac{1}{2}\log|\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{SZD}| - \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{Py}\ . \qquad (5.11)$$

Using $\underset{\sim}{R} = \gamma_0\underset{\sim}{I}_N$ and (2.52) this becomes

$$L_1 = -\tfrac{1}{2}(N - p^*)\log\gamma_0 - \tfrac{1}{2}\log|\underset{\sim}{X}^{*\prime}\underset{\sim}{X}^{*}| - \tfrac{1}{2}\log|\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{MZD}/\sigma_0^2| - \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{Py} \qquad (5.12)$$

and then, because $\underset{\sim}{D}_0 = \underset{\sim}{D}/\sigma_0^2$ as in (2.55), this is also

$$L_1 = -\tfrac{1}{2}(N - p^*)\log\gamma_0 - \tfrac{1}{2}\log|\underset{\sim}{X}^{*\prime}\underset{\sim}{X}^{*}| - \tfrac{1}{2}\log|\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{MZD}_0| - \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{Py}\ . \qquad (5.13)$$

Since $\underset{\sim}{D}_0$ is a function of $\gamma_1, \cdots, \gamma_c$ but not of γ_0, only the first and last terms of (5.13) involve γ_0 .

5.3. THE REML EQUATIONS

a. General results

We will use θ_i to represent either σ_i^2 or γ_i. Then for differentiating L_1 we utilize (2.4) and (2.80) to get

$$\begin{aligned}
\frac{\partial L_1}{\partial\theta_i} &= -\tfrac{1}{2}\,\mathrm{tr}\left(\underset{\sim}{V}^{-1}\frac{\partial\underset{\sim}{V}}{\partial\theta_i}\right) + \tfrac{1}{2}\,\mathrm{tr}\left[(\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^{*})^{-1}\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\frac{\partial\underset{\sim}{V}}{\partial\theta_i}\underset{\sim}{V}^{-1}\underset{\sim}{X}^{*}\right] + \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{P}\frac{\partial\underset{\sim}{V}}{\partial\theta_i}\underset{\sim}{Py} \\
&= -\tfrac{1}{2}\,\mathrm{tr}\left\{[\underset{\sim}{V}^{-1} - \underset{\sim}{V}^{-1}\underset{\sim}{X}^{*}(\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^{*})^{-1}\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}]\frac{\partial\underset{\sim}{V}}{\partial\theta_i}\right\} + \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{P}\frac{\partial\underset{\sim}{V}}{\partial\theta_i}\underset{\sim}{Py} \\
&= -\tfrac{1}{2}\,\mathrm{tr}\left(\underset{\sim}{P}\frac{\partial\underset{\sim}{V}}{\partial\theta_i}\right) + \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{P}\frac{\partial\underset{\sim}{V}}{\partial\theta_i}\underset{\sim}{Py}\ , \qquad (5.14)
\end{aligned}$$

on using (2.26) and (2.59); and with (3.23) this is also

$$\frac{\partial L_1}{\partial\theta_i} = -\tfrac{1}{2}\,\mathrm{tr}\left(\underset{\sim}{P}\frac{\partial\underset{\sim}{V}}{\partial\theta_i}\right) + \tfrac{1}{2}(\underset{\sim}{y} - \underset{\sim}{X}\hat{\underset{\sim}{\alpha}})'\underset{\sim}{V}^{-1}\frac{\partial\underset{\sim}{V}}{\partial\theta_i}\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\hat{\underset{\sim}{\alpha}})\ , \qquad (5.15)$$

a result found as the first equation in [H, Sec. 5].

Equating (5.14) to $\underset{\sim}{0}$, and denoting solutions by, for example, $\hat{\theta}$ to distinguish them from $\tilde{\theta}$ of ML, gives

$$\operatorname{tr}\left(\hat{\underset{\sim}{P}} \frac{\partial \hat{\underset{\sim}{V}}}{\partial \theta_i}\right) = \underset{\sim}{y}'\hat{\underset{\sim}{P}} \frac{\partial \hat{\underset{\sim}{V}}}{\partial \theta_i} \hat{\underset{\sim}{P}}\underset{\sim}{y} \quad \text{for } i = 0, \cdots, c. \tag{5.16}$$

b. Equations for $\hat{\underset{\sim}{\sigma}}^2$

(i) The main result Using $\theta_i = \sigma_i^2$ for $i = 0, \cdots, c$, equations (5.16) are

$$\operatorname{tr}(\hat{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') = \underset{\sim}{y}'\hat{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\hat{\underset{\sim}{P}}\underset{\sim}{y} \, , \tag{5.17}$$

$$= (\underset{\sim}{y} - \underset{\sim}{X}\hat{\hat{\underset{\sim}{\alpha}}})'\hat{\underset{\sim}{V}}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\hat{\underset{\sim}{V}}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\hat{\hat{\underset{\sim}{\alpha}}}) \tag{5.18}$$

where $(\underset{\sim}{y} - \underset{\sim}{X}\hat{\hat{\underset{\sim}{\alpha}}}) = \hat{\underset{\sim}{V}}\hat{\underset{\sim}{P}}\underset{\sim}{y}$, in accord with (3.22). The symbol $\hat{\hat{\underset{\sim}{\alpha}}}$ is used to contrast both with $\hat{\underset{\sim}{\alpha}}$ of the generalized least squares equations (3.1), and with $\tilde{\underset{\sim}{\alpha}}$ of the ML equations (4.5). $\tilde{\alpha}$ is $\hat{\alpha}$ with V replaced by $\tilde{\underset{\sim}{V}}$ and $\hat{\hat{\alpha}}$ is $\hat{\alpha}$ with V replaced by $\hat{\underset{\sim}{V}}$.

Equations (5.17) can be expressed in a variety of ways. First, the left-hand side can utilize the identity

$$\operatorname{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') = \operatorname{tr}(\underset{\sim}{P}\underset{\sim}{V}\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') = \operatorname{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{V})$$

$$= \operatorname{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P} \sum_{j=0}^{c} \sigma_j^2\underset{\sim}{Z}_j\underset{\sim}{Z}_j') = \sum_{j=0}^{c} \operatorname{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\sigma_j^2 \, ,$$

so enabling (5.17) to be expressed as

$$\left\{\operatorname{tr}(\hat{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\hat{\underset{\sim}{P}}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\right\}\hat{\underset{\sim}{\sigma}}^2 = \left\{\underset{\sim}{y}'\hat{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\hat{\underset{\sim}{P}}\underset{\sim}{y}\right\}, \quad \text{for } i,j = 0, \cdots, c. \tag{5.19}$$

Equations (5.17), (5.18) or (5.19) can be considered the main REML equations. Non-negative solutions are REML estimators.

(ii) A single equation for σ_0^2 Separating out the case of $i = 0$ from (5.17) is interesting. It gives

$$\operatorname{tr}(\hat{\underset{\sim}{P}}) = \underset{\sim}{y}'\hat{\underset{\sim}{P}}^2\underset{\sim}{y} \, . \tag{5.20}$$

To get this in a more attractive form we use the same procedure as in Sec. 4.3c. Multiply (5.17) by $\hat{\sigma}_i^2$ and sum over $i = 1, \cdots, c$:

$$\mathrm{tr}\left(\hat{\mathbf{P}} \sum_{i=1}^{c} \hat{\sigma}_i^2 \mathbf{Z}_i \mathbf{Z}_i'\right) = \mathbf{y}'\hat{\mathbf{P}} \sum_{i=1}^{c} \hat{\sigma}_i^2 \mathbf{Z}_i \mathbf{Z}_i' \hat{\mathbf{P}}\mathbf{y}$$

Now use $\sum_{i=1}^{c} \hat{\sigma}_i^2 \mathbf{Z}_i \mathbf{Z}_i' = \hat{\mathbf{V}} - \hat{\sigma}_0 \mathbf{I}$ and so get

$$\mathrm{tr}(\hat{\mathbf{P}}\hat{\mathbf{V}}) - \hat{\sigma}_0^2 \mathrm{tr}(\hat{\mathbf{P}}) = \mathbf{y}'\hat{\mathbf{P}}\hat{\mathbf{V}}\hat{\mathbf{P}}\mathbf{y} - \hat{\sigma}_0^2 \mathbf{y}'\hat{\mathbf{P}}^2\mathbf{y} .$$

Using (2.31), (5.20) and $\hat{\mathbf{P}}\hat{\mathbf{V}}\hat{\mathbf{P}} = \hat{\mathbf{P}}$ along with (3.32) reduces this to

$$N - p^* = (\mathbf{y} - \mathbf{X}\hat{\hat{\boldsymbol{\alpha}}})'\hat{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\hat{\boldsymbol{\alpha}}}) \tag{5.21}$$

which, along with $\mathbf{V} = \sigma_0^2 \mathbf{H}$, gives

$$\hat{\sigma}_0^2 = \frac{(\mathbf{y} - \mathbf{X}\hat{\hat{\boldsymbol{\alpha}}})'\hat{\mathbf{H}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\hat{\boldsymbol{\alpha}}})}{N - p^*} . \tag{5.22}$$

This and (5.17) or (5.18) for $i = 1, \cdots, c$ then constitute a set of REML equations:

$$\mathrm{tr}(\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i') = \mathbf{y}'\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\mathbf{y}$$

$$= (\mathbf{y} - \mathbf{X}\hat{\hat{\boldsymbol{\alpha}}})'\hat{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{V}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\hat{\boldsymbol{\alpha}}}) \tag{5.23}$$

$$= (\mathbf{y} - \mathbf{X}\hat{\hat{\boldsymbol{\alpha}}})'\hat{\mathbf{H}}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{H}}^{-1}(\mathbf{y} - \mathbf{X}\hat{\hat{\boldsymbol{\alpha}}})/\hat{\sigma}_0^4 \tag{5.24}$$

for $i = 1, \cdots, c$.

(iii) Derivation from ML equations   Another derivation of (5.17) is based on the ML equations (4.6):

$$\mathrm{tr}(\tilde{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i') = \mathbf{y}'\tilde{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\tilde{\mathbf{P}}\mathbf{y}, \quad \text{for } i = 0, \cdots, 1, \tag{5.25}$$

which come from applying ML to

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\alpha} + \underset{\sim}{Z}\underset{\sim}{b} + \underset{\sim}{e} \sim \mathcal{N}(\underset{\sim}{X}\underset{\sim}{\alpha},\ \underset{\sim}{V}) . \tag{5.26}$$

REML is just ML applied to

$$\underset{\sim}{K}'\underset{\sim}{y} = \underset{\sim}{K}'\underset{\sim}{Z}\underset{\sim}{b} + \underset{\sim}{K}'\underset{\sim}{e} \sim \mathcal{N}(\underset{\sim}{0},\ \underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K}) . \tag{5.27}$$

The differences between (5.26) and (5.27) are that (5.27) is just (5.26) with $\underset{\sim}{y}$, $\underset{\sim}{X}$, $\underset{\sim}{Z}$ and $\underset{\sim}{V}$ replaced by $\underset{\sim}{K}'\underset{\sim}{y}$, $\underset{\sim}{0}$, $\underset{\sim}{K}'\underset{\sim}{Z}$ and $\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K}$, respectively. Making these replacements in (5.25) therefore yields the REML equations; in carrying this through, note that the replacements have the effect of making $\underset{\sim}{P}$ become $(\underset{\sim}{K}\underset{\sim}{V}\underset{\sim}{K}')^{-1}$, because $\underset{\sim}{K}'\underset{\sim}{X} = \underset{\sim}{0}$. Hence the REML equations are

$$\mathrm{tr}[(\underset{\sim}{K}'\hat{\underset{\sim}{V}}\underset{\sim}{K})^{-1}\underset{\sim}{K}'\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{K}] = \underset{\sim}{y}'\underset{\sim}{K}(\underset{\sim}{K}'\hat{\underset{\sim}{V}}\underset{\sim}{K})^{-1}\underset{\sim}{K}'\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{K}(\underset{\sim}{K}'\hat{\underset{\sim}{V}}\underset{\sim}{K})^{-1}\underset{\sim}{K}'\underset{\sim}{y}$$

and on using $\underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K} = \underset{\sim}{P}$ of (2.75) this becomes

$$\mathrm{tr}(\hat{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') = \underset{\sim}{y}'\hat{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\hat{\underset{\sim}{P}}\underset{\sim}{y}, \quad \text{for } i = 0, \cdots, 1,$$

which is precisely (5.17).

c. Equations for $\dot{\underset{\sim}{\gamma}}$

Using $\theta_i = \gamma_0$ and $\partial\underset{\sim}{V}/\partial\gamma_0 = \underset{\sim}{H}$ from (2.84) gives (5.15) as

$$\begin{aligned} \frac{\partial L_1}{\partial \gamma_0} &= -\tfrac{1}{2}\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{H}) + \tfrac{1}{2}(\underset{\sim}{y} - \underset{\sim}{X}\hat{\underset{\sim}{\alpha}})\underset{\sim}{V}^{-1}\underset{\sim}{H}\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\hat{\underset{\sim}{\alpha}}) \\ &= -\tfrac{1}{2}\gamma_0^{-1}[N - p^* - \underset{\sim}{y}'\underset{\sim}{S}(\underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{b}})] , \end{aligned} \tag{5.28}$$

after using $\gamma_0\underset{\sim}{H} = \underset{\sim}{V}$, $\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{V}) = N - p^*$ and (3.33). This is expression [H5.5]. In equating it to 0 and using (3.33) again we get

$$N - p^* = (\underset{\sim}{y} - \underset{\sim}{X}\hat{\hat{\underset{\sim}{\alpha}}})'\hat{\underset{\sim}{V}}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\hat{\hat{\underset{\sim}{\alpha}}}) \tag{5.29}$$

which is (5.21) and so leads to (5.22).

In using $\theta_i = \gamma_i$ for i = 1, ⋯, c in (5.14), to derive first a result given in Harville [1977], we note that with

$$\underset{\sim}{D} = \text{diag}\{\sigma_1^2 \underset{\sim}{I}_{q_1} \; \cdots \; \sigma_i^2 \underset{\sim}{I}_{q_i} \; \cdots \; \sigma_c^2 \underset{\sim}{I}_{q_c}\}$$

of (1.9)

$$\frac{\partial \underset{\sim}{D}}{\partial \sigma_i^2} = \underset{\sim}{\Delta}_i$$

for $\underset{\sim}{\Delta}_i$ defined in (4.19). Then, for (5.14) for i = 1, ⋯, c, starting at (2.86),

$$\frac{\partial \underset{\sim}{V}}{\partial \gamma_i} = \gamma_0 \underset{\sim}{Z}_i \underset{\sim}{Z}_i' = \sigma_0^2 \underset{\sim}{Z} \underset{\sim}{\Delta}_i \underset{\sim}{Z}' = \underset{\sim}{Z} \underset{\sim}{\Delta}_i \sigma_0^2 \underset{\sim}{Z}' = \underset{\sim}{Z} \frac{\partial \underset{\sim}{D}}{\partial \sigma_i^2} \frac{\partial \sigma_i^2}{\partial \gamma_i} \underset{\sim}{Z}' = \underset{\sim}{Z} \frac{\partial \underset{\sim}{D}}{\partial \gamma_i} \underset{\sim}{Z}' \; . \qquad (5.30)$$

Using this, (5.14) is

$$\frac{\partial L_1}{\partial \gamma_i} = -\tfrac{1}{2}\text{tr}\left(\underset{\sim}{Z}'\underset{\sim}{P}\underset{\sim}{Z} \frac{\partial \underset{\sim}{D}}{\partial \gamma_i}\right) + \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z} \frac{\partial \underset{\sim}{D}}{\partial \gamma_i} \underset{\sim}{Z}'\underset{\sim}{P}\underset{\sim}{y} \; , \qquad (5.31)$$

which, from (3.13) and (3.25) is

$$\frac{\partial L_1}{\partial \gamma_i} = -\tfrac{1}{2}\text{tr}\left[(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{S}\underset{\sim}{Z}\underset{\sim}{D})^{-1}\underset{\sim}{Z}'\underset{\sim}{S}\underset{\sim}{Z} \frac{\partial \underset{\sim}{D}}{\partial \gamma_i}\right] + \tfrac{1}{2}\hat{\underset{\sim}{\nu}}' \frac{\partial \underset{\sim}{D}}{\partial \gamma_i} \hat{\underset{\sim}{\nu}} \; . \qquad (5.32)$$

This is the third equation in the right-hand column of page 326 of Harville [1977].

Of course, using $\partial \underset{\sim}{V}/\partial \gamma_i = \gamma_0 \underset{\sim}{Z}_i \underset{\sim}{Z}_i'$ of (2.86) directly in (5.16) gives

$$\text{tr}(\hat{\underset{\sim}{P}}\hat{\gamma}_0 \underset{\sim}{Z}_i \underset{\sim}{Z}_i') = \underset{\sim}{y}'\hat{\underset{\sim}{P}}\hat{\gamma}_0 \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \hat{\underset{\sim}{P}}\underset{\sim}{y}$$

or

$$\text{tr}(\hat{\underset{\sim}{P}}\underset{\sim}{Z}_i \underset{\sim}{Z}_i') = \underset{\sim}{y}'\hat{\underset{\sim}{P}}\underset{\sim}{Z}_i \underset{\sim}{Z}_i' \hat{\underset{\sim}{P}}\underset{\sim}{y} \quad \text{for } i = 1, \cdots, c$$

as the equations to be used in conjunction with (5.22). These are the same as (5.23) and (5.24).

d. Comparisons with ML

From (4.63) and (5.19) we have the following.

$$\text{ML:}\quad \{\text{tr}(\tilde{V}^{-1}Z_iZ_i'\tilde{V}^{-1}Z_jZ_j')\}\tilde{\sigma}^2 = \{y'\tilde{P}Z_iZ_i'\tilde{P}y\} \qquad (4.63)$$

$$\text{REML:}\quad \{\text{tr}(\hat{P}Z_iZ_i'\hat{P}Z_jZ_j')\}\hat{\sigma}^2 = \{y'\hat{P}Z_iZ_i'\hat{P}y\} \qquad (5.19)$$

for $i = 0, 1, \cdots, c$.

Clearly, where the ML equations have $\tilde{V}^{-1}$ in their left-hand side, the REML equations have $\hat{P}$. This is the only difference between the two sets of equations. We also have

$$\text{ML:}\quad \tilde{\sigma}_0^2 = (y - X\tilde{\alpha})'H^{-1}(y - X\tilde{\alpha})/N \qquad (4.12)$$

$$\text{REML:}\quad \hat{\sigma}_0^2 = (y - X\hat{\hat{\alpha}})'\hat{H}^{-1}(y - X\hat{\hat{\alpha}})/(N - p^*) \qquad (5.22)$$

Here we see in contrast to ML, that in REML the p^* degrees of freedom for the fixed effects are taken into account in estimating σ_0^2. Used in conjunction with (4.12) and (5.22) respectively are

$$\text{ML:}\quad \text{tr}(\tilde{H}^{-1}Z_iZ_i') = (y - X\tilde{\alpha})'\tilde{H}^{-1}Z_iZ_i'\tilde{H}^{-1}(y - X\tilde{\alpha})/\tilde{\sigma}_0^2 \qquad (4.13)$$

$$\text{REML:}\quad \text{tr}(\hat{P}Z_iZ_i') = (y - X\hat{\hat{\alpha}})'\hat{H}^{-1}Z_iZ_i'\hat{H}^{-1}(y - X\hat{\hat{\alpha}})/\hat{\sigma}_0^4 . \qquad (5.24)$$

It is noticeable here that the left-hand side of (5.24) has $\hat{P}$ where that of (4.13) has $\tilde{H}^{-1}$; and for the same reason the denominator of the right-hand side of (5.24) is $\hat{\sigma}_0^4$ whereas that of (4.13) is $\tilde{\sigma}_0^2$.

e. Comparisons with MME's

Just as with ML, in Section 4.3d, we now show how the REML equations (5.22) and (5.23) can be expressed in terms of the MME's. First, in the same way that (4.17) was derived from (4.13) for the ML estimator, so also does (5.22), which differs from (4.13) only in its denominator, become

$$\hat{\sigma}_0^2 = \frac{\mathbf{y}'(\mathbf{y} - \mathbf{X}\hat{\hat{\alpha}} - \mathbf{Z}\hat{\hat{\mathbf{b}}})}{N - p^*} \tag{5.33}$$

where $\hat{\hat{\mathbf{b}}}$, like $\hat{\hat{\alpha}}$, is defined in terms of $\tilde{\mathbf{b}}$ the MME's (3.3) or, equivalently (3.4) and (3.5) — in either case, using the REML $\hat{\mathbf{V}}$ in those equations in place of $\mathbf{V}$.

Equation (5.23) for i = 1, ⋯, c is

$$\mathrm{tr}(\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i') = \mathbf{y}'\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\mathbf{y} \tag{5.23}$$

similar to (4.18). The left-hand side here is the same as that of (4.18) but with $\tilde{\mathbf{V}}^{-1}$ replaced by $\hat{\mathbf{P}}$. Now, from (2.14) and (2.46) respectively,

$$\mathbf{V}^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}\mathbf{D}(\mathbf{I} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D})^{-1}\mathbf{Z}'\mathbf{R}^{-1} \quad \text{and} \quad \mathbf{P} = \mathbf{S} - \mathbf{S}\mathbf{Z}\mathbf{D}(\mathbf{I} + \mathbf{Z}'\mathbf{S}\mathbf{Z}\mathbf{D})^{-1}\mathbf{Z}'\mathbf{S} \; .$$

Notice that in $\mathbf{V}^{-1}$, replacing $\mathbf{R}^{-1}$ by $\mathbf{S}$ yields $\mathbf{P}$; and the same replacement in $\mathbf{T}^* = (\mathbf{I} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D})^{-1}$ yields $\mathbf{T} = (\mathbf{I} + \mathbf{Z}'\mathbf{S}\mathbf{Z}\mathbf{D})^{-1}$. Therefore (4.23) for the left-hand side of (4.18), with $\mathbf{R}^{-1}$ replaced by $\mathbf{S}$ applies to the left-hand side of (5.23). This gives

$$\mathrm{tr}(\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i') = [q_i - \mathrm{tr}(\mathbf{T}_{ii})]/\sigma_i^2 \tag{5.34}$$

for

$$\mathbf{T} = (\mathbf{I} + \mathbf{Z}'\mathbf{S}\mathbf{Z}\mathbf{D})^{-1} = \begin{bmatrix} \mathbf{T}_{11} & \cdots & \mathbf{T}_{1c} \\ \vdots & & \\ \mathbf{T}_{c1} & \cdots & \mathbf{T}_{cc} \end{bmatrix} . \tag{5.35}$$

Furthermore, the right-hand side of (5.23) is identical to that of (4.18) for which (4.24) applies. The outcome of this is equations just like (4.26) and (4.27) results — with $\mathbf{T}_{ii}$ in place of $\mathbf{T}^*_{ii}$. Thus the iterative procedure for REML can be expressed as

$$\hat{\sigma}_0^{2\,(r+1)} = \frac{\underset{\sim}{y}'[\underset{\sim}{y} - \underset{\sim}{X}\hat{\hat{\underset{\sim}{\alpha}}}^{(r)} - \underset{\sim}{Z}\hat{\hat{\underset{\sim}{b}}}^{(r)}]}{N - p^*} \tag{5.36}$$

and, for i = 1, ···, c,

$$\hat{\sigma}_i^{2\,(r+1)} = \frac{\hat{\hat{\underset{\sim}{b}}}_i'^{(r)}\hat{\hat{\underset{\sim}{b}}}_i^{(r)} + \hat{\sigma}_i^{2\,(r)}\mathrm{tr}(\underset{\sim}{T}_{ii}^{(r)})}{q_i} \tag{5.37}$$

or

$$\hat{\sigma}_i^{2\,(r+1)} = \frac{\hat{\hat{\underset{\sim}{b}}}_i'^{(r)}\hat{\hat{\underset{\sim}{b}}}_i^{(r)}}{q_i - \mathrm{tr}(\underset{\sim}{T}_{ii}^{(r)})} . \tag{5.38}$$

Again, as in (4.26) and (4.27), this iteration procedure always yields positive estimates. The second part of Harville's lemma ensures this:

Lemma: (i) $\mathrm{tr}(\underset{\sim}{T}_{ii})$ is positive;

(ii) $q_i \geq \mathrm{tr}(\underset{\sim}{T}_{ii})$ for $\sigma_i^2 > 0$ with strict inequality holding if $r(\underset{\sim}{X}\ \ \underset{\sim}{Z}_i) > p^*$.

Proof: (i) is proven in exactly the same manner as is (i) of the lemma following (4.27). And for (ii), equation (5.34) gives

$$\begin{aligned}\mathrm{tr}[q_i - \mathrm{tr}(\underset{\sim}{T}_{ii})]/\sigma_i^2 &= \mathrm{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i')\\ &= \mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_i)\\ &= \mathrm{var}(\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}),\ \because\ \underset{\sim}{P}\underset{\sim}{V}\underset{\sim}{P} = \underset{\sim}{P}\\ &\geq 0 .\end{aligned}$$

Therefore, providing $\sigma_i^2 > 0$, $q_i \geq \mathrm{tr}(\underset{\sim}{T}_{ii})$.

For any $\underset{\sim}{Z}_i$, $r(\underset{\sim}{X} \;\; \underset{\sim}{Z}_i) \geq p^*$; but it is easily shown that $\underset{\sim}{Z}_i'\underset{\sim}{P} = \underset{\sim}{0}$ if and only if $\underset{\sim}{Z}_i = \underset{\sim}{X}\underset{\sim}{Q}$ for some $\underset{\sim}{Q}$, in which case $\text{var}(\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}) = \underset{\sim}{0}$ and $r(\underset{\sim}{X} \;\; \underset{\sim}{Z}_i) = p^*$. Therefore, whenever $r(\underset{\sim}{X} \;\; \underset{\sim}{Z}_i) > p^*$, the strict inequality $q_i > \text{tr}(\underset{\sim}{T}_{ii})$ holds, provided $\sigma_i^2 > 0$.
Q.E.D.

The iterative procedures of (5.36) - (5.38) can also be derived from Harville's result given in (5.32). Obviously this is so because (5.23) and (5.32) are just alternative forms of (5.14) for i = 1, ⋯, c.

5.4. THE INFORMATION MATRIX

Somewhat of a compendium-like presentation is given here, similar to that of Section 4.4.

a. For $\underset{\sim}{\dot\sigma}^2$

The information matrix for $\underset{\sim}{\dot\sigma}^2$ based on $\underset{\sim}{y} \sim \mathcal{N}(\underset{\sim}{X}\underset{\sim}{\alpha}, \underset{\sim}{V})$ is, as in (4.33)

$$\underset{\sim}{I}(\underset{\sim}{\dot\sigma}^2) = \tfrac{1}{2}\left\{\text{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{V}^{-1}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\right\} \quad \text{for } i,j = 0, \cdots, c. \tag{4.33}$$

We denote the information matrix based on $\underset{\sim}{K}'\underset{\sim}{y}$ as $\underset{\sim}{I}(\underset{\sim}{\dot\sigma}^2)^*$, to distinguish it from $\underset{\sim}{I}(\underset{\sim}{\dot\sigma}^2)$. Then, because $\underset{\sim}{K}'\underset{\sim}{y} \sim \mathcal{N}(\underset{\sim}{0}, \underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})$, we can derive $\underset{\sim}{I}(\underset{\sim}{\dot\sigma}^2)^*$ from $\underset{\sim}{I}(\underset{\sim}{\dot\sigma}^2)$ by making in $\underset{\sim}{I}(\underset{\sim}{\dot\sigma}^2)$ the same replacements as were made in Sec. 5.3b(iii) for deriving the REML equations from the ML equations, namely replace $\underset{\sim}{V}$ by $\underset{\sim}{K}\underset{\sim}{V}\underset{\sim}{K}'$ and $\underset{\sim}{Z}$ by $\underset{\sim}{K}'\underset{\sim}{Z}$. Doing this to (4.33) gives

$$\begin{aligned}\underset{\sim}{I}(\underset{\sim}{\dot\sigma}^2)^* &= \tfrac{1}{2}\left\{\text{tr}[(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K}'\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K}'\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{K}]\right\} \\ &= \tfrac{1}{2}\left\{\text{tr}[\underset{\sim}{Z}_i'\underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K}'\underset{\sim}{Z}_j][\underset{\sim}{Z}_j'\underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K}'\underset{\sim}{Z}_i]\right\} \\ &= \tfrac{1}{2}\left\{\text{tr}[\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j(\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j)']\right\} .\end{aligned} \tag{5.39}$$

Note that with $\hat{\underset{\sim}{P}}$ replacing $\underset{\sim}{P}$, (5.39) becomes identical to the matrix on the left-hand side of the REML equations (5.19). These can therefore be written as

$$[\underset{\sim}{I}(\underset{\sim}{\dot{\sigma}^2})^*]\hat{\underset{\sim}{\sigma}}^2 = \tfrac{1}{2}\{\underset{\sim}{y}'\hat{\underset{\sim}{P}}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\hat{\underset{\sim}{P}}\underset{\sim}{y}\} ,$$

similar to (4.64) of ML, after using $\underset{\sim}{P} = \underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{V}\underset{\sim}{K})^{-1}\underset{\sim}{K}'$ of (2.75). Separating out the terms for $\underset{\sim}{Z}_0 = \underset{\sim}{I}$ leads to

$$\underset{\sim}{I}(\underset{\sim}{\dot{\sigma}^2})^* = \frac{1}{2}\begin{bmatrix} \mathrm{tr}(\underset{\sim}{P}^2) & \{\mathrm{tr}(\underset{\sim}{Z}_j'\underset{\sim}{P}^2\underset{\sim}{Z}_j)\} \\ \{\mathrm{tr}(\underset{\sim}{Z}_j'\underset{\sim}{P}^2\underset{\sim}{Z}_j)\} & \{\mathrm{tr}[\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j(\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j)']\} \end{bmatrix} \tag{5.40}$$

for $i, j = 1, \cdots, c$.

To simplify notation, define the terms of (5.40) as

$$w_0 = \mathrm{tr}(\underset{\sim}{P}^2) , \tag{5.41}$$

$$\underset{\sim}{w} = \{\mathrm{tr}(\underset{\sim}{Z}_j'\underset{\sim}{P}^2\underset{\sim}{Z}_j)\} \quad \text{for } j = 1, \cdots, c, \tag{5.42}$$

$$\underset{\sim}{W}_{ij} = \underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j , \tag{5.43}$$

and

$$\underset{\sim}{Y} = \{\mathrm{tr}(\underset{\sim}{W}_{ij}\underset{\sim}{W}_{ji})\} \quad \text{for } i, j = 1, \cdots, c, \tag{5.44}$$

so that

$$\underset{\sim}{I}(\underset{\sim}{\dot{\sigma}^2})^* = \frac{1}{2}\begin{bmatrix} w_0 & \underset{\sim}{w}' \\ \underset{\sim}{w} & \underset{\sim}{Y} \end{bmatrix} . \tag{5.45}$$

The terms w_0 and $\underset{\sim}{w}$ in (5.45) are not attractive computationally, whereas the comparable terms of $\underset{\sim}{I}(\underset{\sim}{\dot{\gamma}})^*$ are. We therefore derive $\underset{\sim}{I}(\underset{\sim}{\dot{\gamma}})^*$ from (5.45), using (4.38) of Lemma 4.3. Then, when $\mathrm{var}(\hat{\underset{\sim}{\sigma}}^2)$ is needed, $\mathrm{var}(\hat{\underset{\sim}{\gamma}}) = [\underset{\sim}{I}(\underset{\sim}{\dot{\gamma}})^*]^{-1}$ is derived, and from this, using (4.46), (4.47) and (4.48) applied to REML estimators $\hat{\sigma}_i^2$ and $\hat{\gamma}_i$, the elements of $\mathrm{var}(\hat{\underset{\sim}{\sigma}}^2)$ can be obtained.

b. For $\dot{\gamma}$

An application of (4.38) gives

$$I(\dot{\gamma})^* = J'I(\dot{\sigma}^2)^*J$$

for J of (4.41) and $I(\dot{\sigma}^2)^*$ of (5.45). Hence

$$I(\dot{\gamma})^* = \begin{bmatrix} 1 & \gamma' \\ 0 & \sigma_0^2 I \end{bmatrix} \begin{bmatrix} w_0 & w' \\ w & Y \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \gamma & \sigma_0^2 I \end{bmatrix}$$

$$= \begin{bmatrix} w_0 + 2\gamma' w + \gamma' Y \gamma & \sigma_0^2(w' + \gamma' Y) \\ \sigma_0^2(w + Y\gamma) & \sigma_0^4 Y \end{bmatrix} . \tag{5.46}$$

Simplification of terms in (5.46) is tedious.

$$\gamma' Y = \left\{\sum_j \gamma_j \mathrm{tr}(W_{ij} W_{ji})\right\}$$

$$= \left\{\sum_j \gamma_j \mathrm{tr}(Z_i' P Z_j Z_j' P Z_i)\right\}$$

$$= \left\{\mathrm{tr}[Z_i' P(\sum_{j=1}^{c} \gamma_j Z_j Z_j') P Z_i]\right\}$$

$$= \left\{\mathrm{tr}[Z_i' P(H - I) P Z_i]\right\}, \quad \text{using (1.22);}$$

and with $H = V/\sigma_0^2$ of (1.21) and $PVP = P$ of (2.29) this becomes

$$\gamma' Y = \left\{\mathrm{tr}(Z_i' P Z_i/\sigma_0^2) - \mathrm{tr}(Z_i' P^2 Z_i)\right\}$$

$$= (1/\sigma_0^2)\{\mathrm{tr}(W_{ii})\}' - w' .$$

Hence for (5.46)

$$\sigma_0^2(w' + \gamma' Y) = \{\mathrm{tr}(W_{ii})\}', \quad \text{for } i = 1, \cdots, c. \tag{5.47}$$

And the leading element of (5.46) is

$$
\begin{aligned}
w_0 + 2\underset{\sim}{\gamma}'\underset{\sim}{w} + \underset{\sim}{\gamma}'\underset{\sim}{Y}\underset{\sim}{\gamma} &= w_0 + \underset{\sim}{w}'\underset{\sim}{\gamma} + (\underset{\sim}{w}' + \underset{\sim}{\gamma}'\underset{\sim}{Y})\underset{\sim}{\gamma} \\
&= w_0 + \left\{\mathrm{tr}(\underset{\sim}{Z}_j'\underset{\sim}{P}^2\underset{\sim}{Z}_j)\right\}'\underset{\sim}{\gamma} + (1/\sigma_0^2)\{\mathrm{tr}(\underset{\sim}{W}_{ii})\}'\underset{\sim}{\gamma}, \text{ from (5.42) and (5.47)} \\
&= w_0 + \mathrm{tr}\left(\underset{\sim}{P}^2 \sum_{j=1}^{c} \underset{\sim}{Z}_j\underset{\sim}{Z}_j'\gamma_j\right) + (1/\sigma_0^2)\mathrm{tr}(\underset{\sim}{P}\sum_i \underset{\sim}{Z}_i\underset{\sim}{Z}_i'\gamma_i), \text{ from (5.43)} \\
&= w_0 + \mathrm{tr}[\underset{\sim}{P}^2(\underset{\sim}{H} - \underset{\sim}{I})] + (1/\sigma_0^2)\mathrm{tr}[\underset{\sim}{P}(\underset{\sim}{H} - \underset{\sim}{I})], \text{ from (1.22)} \\
&= w_0 + (1/\sigma_0^2)\mathrm{tr}(\underset{\sim}{P}) - \mathrm{tr}(\underset{\sim}{P}^2) + (1/\sigma_0^4)\mathrm{tr}(\underset{\sim}{PV}) - (1/\sigma_0^2)\mathrm{tr}(\underset{\sim}{P}), \\
&\qquad\qquad \text{after using (1.21) and (2.29)} \\
&= (1/\sigma_0^4)\mathrm{tr}(\underset{\sim}{PV}) \\
&= (N - p^*)/\sigma_0^4\,, \quad \text{from (2.31).} \qquad (5.48)
\end{aligned}
$$

Substituting (5.44), (5.47) and (5.48) into (5.46) gives

$$
\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}})^* = \frac{1}{2}\begin{bmatrix} (N - p^*)/\sigma_0^4 & \{\mathrm{tr}(\underset{\sim}{W}_{ii})\}' \\ \{\mathrm{tr}(\underset{\sim}{W}_{ii})\} & \sigma_0^4\{\mathrm{tr}(\underset{\sim}{W}_{ij}\underset{\sim}{W}_{ji})\} \end{bmatrix} \qquad (5.49)
$$

$$
= \frac{1}{2}\begin{bmatrix} (N - p^*)/\sigma_0^4 & \{\mathrm{tr}(Z_i'PZ_i)\}' \\ \{\mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_i)\} & \sigma_0^4\{\mathrm{tr}(\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{P}\underset{\sim}{Z}_i)\} \end{bmatrix}. \qquad (5.50)
$$

This expression is equivalent to result (42) of Corbeil and Searle [1976a] except that their $\underset{\sim}{W}_{ii}$ and $\underset{\sim}{W}_{ij}$ are, by their definition, σ_0^2 times the $\underset{\sim}{W}_{ii}$ and $\underset{\sim}{W}_{ij}$ used here.

Comparison of $\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}})^*$ in (5.50) with the ML $\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}})$ in (4.37) is of interest: (5.50) is (4.37) with N replaced by $(N - p^*)$ and with $\underset{\sim}{H}^{-1}/\sigma_0^2$, i.e. $\underset{\sim}{V}^{-1}$ replaced by $\underset{\sim}{P}$. This is, perhaps, no surprise since the same replacement in the ML equations (4.12) and (4.18) yields the REML equations (5.21) and (5.23).

c. Relationships with MME's

(i) For $\dot{\underset{\sim}{\sigma}}^2$ On comparing $\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)^*$ with $\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)$ it will be seen that the REML $\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)^*$ of (5.40) is exactly the ML $\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)$ of (4.49) with $\underset{\sim}{V}^{-1}$ replaced by $\underset{\sim}{P}$. Making this change in (4.60), the MME-related form of (4.49), changes N to N - p^* and $\underset{\sim}{T}^*_{ii}$ to $\underset{\sim}{T}_{ii}$, so giving (5.40) identical to

$$\underset{\sim}{I}(\dot{\underset{\sim}{\sigma}}^2)^* = \frac{1}{2}\begin{bmatrix} \dfrac{N-p^*-q}{\sigma_0^4} + \dfrac{1}{\sigma_0^2}\sum\limits_{i=1}^{c}\sum\limits_{j=1}^{c}\mathrm{tr}(\underset{\sim}{T}_{ij}\underset{\sim}{T}_{ji}) & \left\{[\mathrm{tr}(\underset{\sim}{T}_{ii}) - \sum\limits_{k=1}^{c}\mathrm{tr}(\underset{\sim}{T}_{ik}\underset{\sim}{T}_{ki})]/\sigma_0^2\sigma_i^2\right\} \\ \text{sym.} & \begin{Bmatrix} \text{diagonal submatrices: } \mathrm{tr}(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}_{ii})^2/\sigma_i^4 \\ \text{off-diag. submatrices: } \mathrm{tr}(\underset{\sim}{T}_{ij}\underset{\sim}{T}_{ji})/\sigma_i^2\sigma_j^2 \end{Bmatrix} \end{bmatrix}$$

(5.51)

(ii) For $\dot{\underset{\sim}{\gamma}}$ The same replacement is evident in comparing $\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}})$ of (4.37) with $\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}})^*$ of (5.50), along with also replacing N by N - p^*. Hence making these same replacements in (4.61), the MME-related form of (4.37) gives

$$\underset{\sim}{I}(\dot{\underset{\sim}{\gamma}})^* = \frac{1}{2}\begin{bmatrix} (N - p^*)/\sigma_0^4 & \left\{[q_i - \mathrm{tr}(\underset{\sim}{T}_{ii})]/\sigma_i^2\right\}' \\ \left\{[q_i - \mathrm{tr}(\underset{\sim}{T}_{ii})]/\sigma_i^2\right\} & \begin{Bmatrix} \text{diagonal submatrices: } \mathrm{tr}(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}_{ii})^2/\gamma_i^2 \\ \text{off-diag. submatrices: } \mathrm{tr}(\underset{\sim}{T}_{ij}\underset{\sim}{T}_{ji})/\gamma_i\gamma_j \end{Bmatrix} \end{bmatrix}$$

(5.52)

The terms of (5.52) are the first four expressions given by Harville [1977] following [H5.5], in the form

$$-2E\frac{\partial^2 L_1}{\partial\gamma_i^2} = \mathrm{tr}(\underset{\sim}{I} - \underset{\sim}{T}_{ii})^2/\gamma_i^2$$

$$-2E\frac{\partial^2 L_1}{\partial\gamma_i\partial\gamma_k} = \mathrm{tr}(\underset{\sim}{T}_{ik}\underset{\sim}{T}_{ki})/\gamma_i\gamma_k \quad \text{for } i \neq k$$

$$-2E\frac{\partial^2 L_1}{\partial\gamma_i\partial\gamma_0} = \mathrm{tr}(\underset{\sim}{I} - \underset{\sim}{T}_{ii})/\sigma_i^2 = [q_i - \mathrm{tr}(\underset{\sim}{T}_{ii})]/\gamma_0\gamma_i$$

$$-2E\frac{\partial^2 L_1}{\partial\gamma_0^2} = (N - p^*)\sigma_0^{-4} .$$

The first two of these are $\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i')$, and $\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_k\underset{\sim}{Z}_k')$ for $i \neq k$; they are exactly the same as (4.57) and (4.58) respectively, only with $\underset{\sim}{P}$ in place of $\underset{\sim}{V}^{-1}$, but their reductions to functions of $\underset{\sim}{T}_{ij}$'s proceed exactly along the same lines as (4.57) and (4.58). The last two utilize $\partial\underset{\sim}{V}/\partial\gamma_0 = \underset{\sim}{H}$ of (2.84). Its use is vital. For example, from (5.15)

$$\frac{\partial L_1}{\partial\gamma_0} = -\tfrac{1}{2}\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{H}) + \tfrac{1}{2}\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{H}\underset{\sim}{P}\underset{\sim}{y} = -(N - p^* - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{y})/2\gamma_0$$

so that

$$-2E\frac{\partial^2 L_1}{\partial\gamma_0^2} = E\left[\frac{-(N - p^* - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{y})}{\gamma_0^2} + \frac{\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{H}\underset{\sim}{P}\underset{\sim}{y}}{\gamma_0}\right]$$

$$= E\left[\frac{-(N - p^*) + 2\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{y}}{\gamma_0^2}\right] \quad \because \underset{\sim}{P}\underset{\sim}{H}\underset{\sim}{P} = \underset{\sim}{P}\underset{\sim}{V}\underset{\sim}{P}/\gamma_0 = \underset{\sim}{P}/\gamma_0$$

$$= [-(N - p^*) + 2\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{V})]/\gamma_0^2$$

$$= (N - p^*)/\sigma_0^4 .$$

In carrying out these derivations one would use $\underset{\sim}{T}$ defined in (5.35) and, comparable to (4.21)

$$\underset{\sim}{T}_{ii} + \sigma_i^2 \sum_{j=1}^{c} \underset{\sim}{T}_{ij}\underset{\sim}{Z}_j'\underset{\sim}{S}\underset{\sim}{Z}_i = \underset{\sim}{I}_{q_i} \tag{5.53}$$

and

$$\underset{\sim}{T}_{ik} + \sigma_k^2 \sum_{j=1}^{c} \underset{\sim}{T}_{ij}\underset{\sim}{Z}_j'\underset{\sim}{S}\underset{\sim}{Z}_k = 0 \quad \text{for } i \neq k. \tag{5.54}$$

(iii) Another expression for the REML equations The REML equations have been expressed in terms of the information matrix function $\underset{\sim}{I}(\hat{\underset{\sim}{\sigma}}^2)$ following (5.39). In view of (5.40) for $\underset{\sim}{I}(\underset{\sim}{\sigma}^2)$ this means the REML equations are

$$\begin{bmatrix} \operatorname{tr}(\hat{\underset{\sim}{P}}^2) & \{\operatorname{tr}(\hat{\underset{\sim}{P}}^2 \underset{\sim}{Z}_i \underset{\sim}{Z}_i')\}' \\ \{\operatorname{tr}(\hat{\underset{\sim}{P}}^2 \underset{\sim}{Z}_i \underset{\sim}{Z}_i')\} & \{\operatorname{tr}(\hat{\underset{\sim}{P}} \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \hat{\underset{\sim}{P}} \underset{\sim}{Z}_j \underset{\sim}{Z}_j')\} \end{bmatrix} \begin{bmatrix} \hat{\sigma}_0^2 \\ \hat{\underset{\sim}{\sigma}}^2 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{y}' \hat{\underset{\sim}{P}}^2 \underset{\sim}{y} \\ \underset{\sim}{y}' \hat{\underset{\sim}{P}} \underset{\sim}{Z}_i \underset{\sim}{Z}_i' \hat{\underset{\sim}{P}} \underset{\sim}{y} \end{bmatrix} \tag{5.55}$$

for $i, j = 1, \cdots, c$.

But since (5.40) and (5.51) are equal, all submatrices in (5.55) can be expressed in terms of the MME functions directly from (5.51); and the right-hand sides of the equation are available from (4.68) and (4.69). There, just like (4.70), the REML equations in the form (5.55) can be expressed as

$$\begin{bmatrix} \dfrac{N-p^{*}-q}{\hat{\sigma}_0^4} + \dfrac{1}{\hat{\sigma}_0^4} \sum_{i=1}^{c} \sum_{j=1}^{c} \operatorname{tr}(\hat{\underset{\sim}{T}}_{ij} \hat{\underset{\sim}{T}}_{ji}) & \left\{[\operatorname{tr}(\underset{\sim}{T}_{ii}) - \sum_{k=1}^{c} \operatorname{tr}(\hat{\underset{\sim}{T}}_{ik} \hat{\underset{\sim}{T}}_{ki})]/\hat{\sigma}_0^2 \hat{\sigma}_i^2\right\} \\ \text{sym.} & \begin{Bmatrix} \text{diagonal submatrices: } \operatorname{tr}(\underset{\sim}{I}_{q_i} - \hat{\underset{\sim}{T}}_{ii})^2/\hat{\sigma}_i^4 \\ \text{off-diag. submatrices: } \operatorname{tr}(\hat{\underset{\sim}{T}}_{ij} \hat{\underset{\sim}{T}}_{ji})/\hat{\sigma}^2 \hat{\sigma}^2 \end{Bmatrix} \end{bmatrix} \begin{bmatrix} \hat{\sigma}_0^2 \\ \hat{\underset{\sim}{\sigma}}^2 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} \underset{\sim}{y}'(\underset{\sim}{y} - \underset{\sim}{X}\hat{\hat{\underset{\sim}{\alpha}}} - \underset{\sim}{Z}\hat{\hat{\underset{\sim}{b}}})/\hat{\sigma}_0^4 - \dfrac{1}{\hat{\sigma}_0^2} \sum_{i=1}^{c} \hat{\hat{\underset{\sim}{b}}}_i' \hat{\hat{\underset{\sim}{b}}}_i / \hat{\sigma}_i^2 \\ \{\hat{\hat{\underset{\sim}{b}}}_i' \hat{\hat{\underset{\sim}{b}}}_i / \hat{\sigma}_i^4\} \end{bmatrix} \tag{5.56}$$

for $i, j = 1, \cdots, c$.

5.5. COMPUTING ALGORITHMS

a. Newton-Raphson

The matrix to be inverted for this method, see (4.63), is shown for ML estimation in (4.71). As has just been done for information matrices, the matrix for REML estimation by Newton-Raphson can be derived from (4.94) by replacing N by $N - p^*$ and $\underset{\sim}{V}^{-1}$ by $\underset{\sim}{P}$. The latter replacement changes $\underset{\sim}{T}^*_{ij}$ in (4.94) to $\underset{\sim}{T}_{ij}$ and so, from (4.94) we get

$$\left\{-2\frac{\partial^2 L_1}{\partial\gamma_i\partial\gamma_j}\right\}$$

$$= \begin{bmatrix} \gamma_0^{-2}(-N + p^* + 2\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{y}) & \{\tilde{\underset{\sim}{v}}_i'\tilde{\underset{\sim}{v}}_i\}' \\ \{\tilde{\underset{\sim}{v}}_i'\tilde{\underset{\sim}{v}}_i\} & \left\{\begin{array}{l} \text{diagonal terms: } -\gamma_i^{-2}\text{tr}(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}_{ii})^2 + 2\gamma_0\gamma_i^{-1}\tilde{\underset{\sim}{v}}_i'(\underset{\sim}{I}_{q_i} - \underset{\sim}{T}_{ii})\tilde{\underset{\sim}{v}}_i \\ \text{off-diag. terms: } -\gamma_i^{-1}\gamma_j^{-1}\text{tr}(\underset{\sim}{T}_{ij}\underset{\sim}{T}_{ji}) - 2\gamma_0\gamma_j^{-1}\tilde{\underset{\sim}{v}}_i'\underset{\sim}{T}_{ij}\tilde{\underset{\sim}{v}}_j \end{array}\right\} \end{bmatrix}$$

$$\text{for } i,j = 1, \cdots, c. \qquad (5.57)$$

With $\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{y} = \underset{\sim}{y}'\underset{\sim}{S}(\underset{\sim}{y} - \underset{\sim}{Z}\tilde{\underset{\sim}{b}})$ of (3.33), the terms in (5.57) are exactly those given as the last four expressions following [H5.5].

b. Fisher's scoring method

As seen in (4.95), this uses the information matrix, computing forms of which are available in (5.51) and (5.52).

We can also show that Fisher's scoring method applied to L_1 leads directly to the iterative solution of REML equations (5.19). In (5.19) write

$$\underset{\sim}{F} \equiv \{\text{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\} \quad \text{for } i,j = 0, 1, \cdots, c \qquad (5.58)$$

and

$$\underset{\sim}{u} \equiv \{\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}\} \quad \text{for } i = 0, 1, \cdots, c. \qquad (5.59)$$

Then an iterative procedure for solving (5.19) is

$$\hat{\underset{\sim}{\sigma}}^{2(r+1)} = \underset{\sim}{F}^{-1}\underset{\sim}{u}\Big|_{\underset{\sim}{\dot{\sigma}}^2 = \underset{\sim}{\dot{\sigma}}^{2(r)}} \tag{5.60}$$

where superscript r denotes the r'th iterate. In contrast, from the Fisher scoring procedure of (4.95), using L_1 in place of L,

$$\hat{\underset{\sim}{\sigma}}^{2(r+1)} = \hat{\underset{\sim}{\sigma}}^{2(r)} - E\left(\frac{\partial^2 L_1}{\partial\underset{\sim}{\dot{\sigma}}^2\partial\underset{\sim}{\dot{\sigma}}^{2\prime}}\right)^{-1}\frac{\partial L_1}{\partial\underset{\sim}{\dot{\sigma}}^2}\Bigg|_{\underset{\sim}{\dot{\sigma}}^2 = \hat{\underset{\sim}{\sigma}}^{2(r)}}, \tag{5.61}$$

and in this expression, from (5.39),

$$E\left(\frac{\partial^2 L_1}{\partial\underset{\sim}{\dot{\sigma}}^2\partial\underset{\sim}{\dot{\sigma}}^{2\prime}}\right) = -\tfrac{1}{2}\{\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\} \quad \text{for } i,j = 0, \cdots, c$$

$$= -\tfrac{1}{2}\underset{\sim}{F}, \quad \text{from (5.58);}$$

and from (5.15), with $\partial\underset{\sim}{V}/\partial\sigma_i^2 = \underset{\sim}{Z}_i\underset{\sim}{Z}_i'$,

$$\frac{\partial L_1}{\partial\underset{\sim}{\dot{\sigma}}^2} = -\tfrac{1}{2}\left\{\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}\right\}$$

$$= -\tfrac{1}{2}\left\{\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{V}) - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}\right\} \quad \because \underset{\sim}{P}\underset{\sim}{V}\underset{\sim}{P} = \underset{\sim}{P},$$

$$= -\tfrac{1}{2}\left\{\mathrm{tr}\left(\sum_{j=0}^{c}\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\sigma_j^2\right) - \underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}\right\}$$

$$= -\tfrac{1}{2}\left\{\mathrm{tr}(\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\right\}\underset{\sim}{\dot{\sigma}}^2 + \tfrac{1}{2}\left\{\underset{\sim}{y}'\underset{\sim}{P}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}\underset{\sim}{y}\right\}$$

$$= -\tfrac{1}{2}\underset{\sim}{F}\underset{\sim}{\dot{\sigma}}^2 + \tfrac{1}{2}\underset{\sim}{u}, \quad \text{from (5.58) and (5.59).}$$

Hence in (5.61)

$$\hat{\underset{\sim}{\sigma}}^{2(r+1)} = \hat{\underset{\sim}{\sigma}}^{2(r)} + (\tfrac{1}{2}\underset{\sim}{F})^{-1}(-\tfrac{1}{2}\underset{\sim}{F}\underset{\sim}{\dot{\sigma}}^2 + \tfrac{1}{2}\underset{\sim}{u})\Big|_{\underset{\sim}{\dot{\sigma}}^2 = \hat{\underset{\sim}{\sigma}}^{2(r+1)}}$$

$$= \underset{\sim}{F}^{-1}\underset{\sim}{u}\Big|_{\underset{\sim}{\dot{\sigma}}^2 = \hat{\underset{\sim}{\sigma}}^{2(r+1)}}$$

which is (5.60). This is exactly the iterative use of the MINQUE equation noted by Patterson and Thompson [1974].

5.6. CORBEIL AND SEARLE'S REML

The REML described by Corbeil and Searle [1976a] uses a special kind of matrix $\underset{\sim}{K}'$, which they call $\underset{\sim}{T}$. It is based on writing the fixed effects part of the model as $\underset{\sim}{X}\underset{\sim}{\mu}$ with $\underset{\sim}{\mu}$ being defined as the vector of cell means of those sub-most cells of the fixed effects factors which contain data. This definition of $\underset{\sim}{\mu}$ leads to $\underset{\sim}{X}$ of $\underset{\sim}{X}\underset{\sim}{\mu}$ being a diagonal matrix of $\mathbf{1}$-vectors, one for each sub-most cell containing data, with order equal to the number of observations in the cell, n_t say, for the t'th such cell. When there are k filled cells, $\underset{\sim}{X}'\underset{\sim}{X} = \text{diag}\{n_t\}$ for t = 1, $\cdots$, k and

$$\underset{\sim}{M} = \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}' = \underset{\sim}{I}_N - \text{diag}\{n_t^{-1}\underset{\sim}{J}_{n_t}\} = \text{diag}\{\underset{\sim}{I}_{n_t} - n_t^{-1}\underset{\sim}{J}_{n_t}\}$$

where $\underset{\sim}{J}_{n_t}$ is square of order n_t, all elements unity. Deleting the n_1'th, $(n_1 + n_2)$'th, $\cdots$, $(n_1 + n_2 + \cdots + n_k)$'th rows from $\underset{\sim}{M}$ yields an appropriate $\underset{\sim}{K}'$ (called $\underset{\sim}{T}$ by Corbeil and Searle). Not only does it satisfy the usual properties of a $\underset{\sim}{K}'$ [e.g. (2.72), (2.73) and (2.75)], but it satisfies the additional result $\underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{K})^{-1}\underset{\sim}{K}' = \underset{\sim}{M}$.

One advantage of this procedure is that it is computationally easy. Another is that it avoids all discussion of how to handle interactions in the fixed effects part of the model, a facet of fixed effects models that is always difficult to handle when data are unbalanced. This procedure effectively assumes all possible fixed effects interactions (for the available data) are present and includes them, along with main effects, in a full rank representation — in effect, for the fixed effects part of the model, the μ_{ij}-model that has been discussed by Searle [1971a] and others. A disadvantage of the procedure is that it does not allow one to assume that certain fixed effects interactions do not exist. However, it has been said that seldom is one in such a well-informed position, so that maybe all interactions should be assumed extant, in which case the procedure is suitable.

Cross References to Harville [1977]

Harville [1977] Equ. No. + No. of lines	Here Equation Number(s)	Harville [1977] Equ. No. + No. of lines	Here Equation Number(s)
(2.1)	(1.2)	p. 326, left, line 25	(5.39)
(2.1) + 8	(1.6)	(5.1) - 2	(5.7)
(2.1) + 9	(1.8), (1.12)	(5.1)	(5.8)
(2.2) + 1	(1.7), (1.10)	(5.2)	(3.30), (3.28)
(2.2) + 2	(1.18)	p. 326, right, line 2	(5.9)
(3.1)	(3.1)	" " " 4	(5.10)
(3.1) + 1	(3.5), (3.21), (3.25)	" " " 10	(5.32)
(3.1) + 2	(2.26)	(5.4) - 4	(2.47), (3.8), (5.35)
(3.3)	(3.4)	(5.4) - 2	(5.53)
(3.4)	(3.4)	(5.4) - 1	(5.54)
(3.5)	(3.26)	(5.5)	(5.28)
(3.5) + 1	(2.36)	(5.5) + 1-4	(5.52)
(3.6)	(2.14)	(5.5) + 5-11	(5.57)
(3.6) + 1	(2.15), (3.12)	p. 328, left, line 20	(4.64)
(3.6) + 2	(2.46), (2.48)	(6.1)	(4.24a), (4.62)
(3.7)	(3.13)	(6.2)	(4.17), (4.25)
(3.8)	(3.3)	(6.2) + 2	(2.16), (3.6)
(3.8) + 38	(4.1)	(6.3)	(5.37)
(4.1)	(4.12)	(6.4)	(5.36)
(4.3) + 21	Below (2.71)	(6.5)	(4.24b), (4.27)
(4.3) + 24	(2.73)	(6.6)	(5.38)
(4.3) + 33	(5.3)	(7.1) - 1	(5.23)
p. 326, left, line 6	(5.15)	(7.1)	(5.23) + (3.30)
" " " 12	(5.39)	(7.2)	(5.23) + (3.28)
" " " 23	(4.33)		

Chapter 6

MINIMUM NORM AND MINIMUM VARIANCE QUADRATIC UNBIASED ESTIMATORS (MINQUE and MIVQUE)

Minimum norm quadratic unbiased estimators (MINQUE) were first developed by Rao [1971b]; he also considered minimum variance quadratic unbiased estimators (MIVQUE), extending the first development of these made by Townsend [1968] and Townsend and Searle [1971]. We give here a general development based on Rao's work, using the same model as previously, namely $\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \ddot{\mathbf{Z}}\dot{\mathbf{b}}$ where $\boldsymbol{\alpha}$ represents fixed effects and $\dot{\mathbf{b}}$ the random effects including $\mathbf{e}$. The approach is to consider the problem of estimating a linear function $\mathbf{p}'\boldsymbol{\sigma}^2$ of the variance components (for known $\mathbf{p}'$) using a quadratic function $\mathbf{y}'\mathbf{A}\mathbf{y}$ of the observations where, without loss of generality, $\mathbf{A}$ is symmetric, $\mathbf{A} = \mathbf{A}'$.

6.1. TRANSLATION INVARIANCE

A quadratic form $\mathbf{y}'\mathbf{A}\mathbf{y}$ used as an estimator of $\mathbf{p}'\boldsymbol{\sigma}^2$ is said to be translation invariant if it is not affected by changes in the fixed effects $\boldsymbol{\alpha}$. This means that if $\boldsymbol{\alpha}$ becomes $\boldsymbol{\alpha} + \boldsymbol{\delta}$ then $\mathbf{y}'\mathbf{A}\mathbf{y}$ is to be unchanged; i.e.,

$$\mathbf{y}'\mathbf{A}\mathbf{y} = (\mathbf{X}\boldsymbol{\alpha} + \ddot{\mathbf{Z}}\dot{\mathbf{b}})'\mathbf{A}(\mathbf{X}\boldsymbol{\alpha} + \ddot{\mathbf{Z}}\dot{\mathbf{b}}) = (\mathbf{X}\boldsymbol{\alpha} + \mathbf{X}\boldsymbol{\delta} + \ddot{\mathbf{Z}}\dot{\mathbf{b}})'\mathbf{A}(\mathbf{X}\boldsymbol{\alpha} + \mathbf{X}\boldsymbol{\delta} + \ddot{\mathbf{Z}}\dot{\mathbf{b}}) ,$$

so giving

$$\boldsymbol{\delta}'\mathbf{X}'\mathbf{A}\mathbf{X}\boldsymbol{\delta} + 2\boldsymbol{\delta}'\mathbf{X}'\mathbf{A}(\mathbf{X}\boldsymbol{\alpha} + \ddot{\mathbf{Z}}\dot{\mathbf{b}}) = \mathbf{0} .$$

We want this to be true for all $\boldsymbol{\delta}$. A sufficient condition for this is to choose $\mathbf{A}$ so that

$$\mathbf{A}\mathbf{X} = \mathbf{0} . \tag{6.1}$$

When this is satisfied, $\mathbf{y}'\mathbf{A}\mathbf{y}$ is said to be a translation invariant estimator.

Translation invariance is a property of both ML and REML estimators (it is a by-product of those estimation procedures), and it is a property that will be imposed upon the MINQUE and MIVQUE procedures.

6.2. UNBIASEDNESS

If $y'Ay$ is to be an unbiased estimator of $p'\dot{\sigma}^2$ then

$$p'\dot{\sigma}^2 = E(y'Ay) = \mathrm{tr}(AV) + \alpha'X'AX\alpha$$

$$= \mathrm{tr}(AV), \text{ from } (6.1) = \mathrm{tr}\left(A \sum_{i=0}^{c} \sigma_i^2 Z_i Z_i'\right) ;$$

i.e.,

$$\sum_{i=0}^{c} p_i \sigma_i^2 = \sum_{i=0}^{c} \mathrm{tr}(AZ_i Z_i')\sigma_i^2 .$$

Requiring this to be true for all values of σ_i^2 implies

$$p_i = \mathrm{tr}(AZ_i Z_i') . \tag{6.2}$$

6.3. MINIMUM NORM ESTIMATION (MINQUE)

If the b_i's in the model were known, Rao suggests that a "natural" estimator of $p'\sigma^2 = \Sigma p_i \sigma_i^2$ would be

$$p'\tilde{\dot{\sigma}}^2 = \Sigma p_i \tilde{\dot{\sigma}}_i^2 = \sum_{i=0}^{c} p_i \sum_{j=1}^{q_i} b_{ij}^2/q_i = \sum_{i=0}^{c} \frac{p_i}{q_i} b_i' b_i . \tag{6.3}$$

Define

$$\dot{\Delta} = \mathrm{diag}\left\{\frac{p_0}{q_0} I_{q_0} \cdots \frac{p_c}{q_c} I_{q_c}\right\} , \tag{6.4}$$

so that (6.3) can be expressed as

$$p'\tilde{\dot{\sigma}}^2 = \dot{b}'\dot{\Delta}\dot{b} . \tag{6.5}$$

In contrast to this estimator, we will use $\underset{\sim}{y}'\underset{\sim}{A}\underset{\sim}{y}$ with $\underset{\sim}{A}\underset{\sim}{X} = \underset{\sim}{0}$, i.e.,

$$\underset{\sim}{p}'\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{y}'\underset{\sim}{A}\underset{\sim}{y} = \dot{\underset{\sim}{b}}'\dot{\underset{\sim}{Z}}'\underset{\sim}{A}\dot{\underset{\sim}{Z}}\dot{\underset{\sim}{b}} , \tag{6.6}$$

with the difference between the estimators being

$$\underset{\sim}{p}'\hat{\underset{\sim}{\sigma}}^2 - \underset{\sim}{p}'\tilde{\underset{\sim}{\sigma}}^2 = \dot{\underset{\sim}{b}}'(\dot{\underset{\sim}{Z}}'\underset{\sim}{A}\dot{\underset{\sim}{Z}} - \dot{\underset{\sim}{\Delta}})\dot{\underset{\sim}{b}} . \tag{6.7}$$

Rao appeals to intuition in calling (6.3) a "natural" estimator, and proceeds to minimize the difference between it and $\underset{\sim}{y}'\underset{\sim}{A}\underset{\sim}{y}$, namely (6.7), by choosing to minimize a Euclidian norm of the matrix $(\dot{\underset{\sim}{Z}}'\underset{\sim}{A}\dot{\underset{\sim}{Z}} - \dot{\underset{\sim}{\Delta}})$ in (6.7). More generally, he minimizes a weighted norm, using weights w_i associated with $\underset{\sim}{b}_i$ for $i = 0, \cdots, p$, represented by

$$\dot{\underset{\sim}{D}}_w = \text{diag}\{w_0\underset{\sim}{I}_{q_0} \cdots w_c\underset{\sim}{I}_{q_c}\}, \text{ with } \dot{\underset{\sim}{V}}_w = \dot{\underset{\sim}{Z}}\dot{\underset{\sim}{D}}_w\dot{\underset{\sim}{Z}}' \text{ similar to (1.30).} \tag{6.8}$$

Then the norm to be minimized is

$$\begin{aligned}
&||\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}(\dot{\underset{\sim}{Z}}'\underset{\sim}{A}\dot{\underset{\sim}{Z}} - \dot{\underset{\sim}{\Delta}})\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}|| \\
&= \text{tr}[\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}(\dot{\underset{\sim}{Z}}'\underset{\sim}{A}\dot{\underset{\sim}{Z}} - \dot{\underset{\sim}{\Delta}})\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}]^2 \\
&= \text{tr}(\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}\dot{\underset{\sim}{Z}}'\underset{\sim}{A}\dot{\underset{\sim}{Z}}\dot{\underset{\sim}{D}}_w\dot{\underset{\sim}{Z}}'\underset{\sim}{A}\dot{\underset{\sim}{Z}}\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}) + \text{tr}(\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}\dot{\underset{\sim}{\Delta}}\dot{\underset{\sim}{D}}_w\dot{\underset{\sim}{\Delta}}\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}) - 2\text{tr}(\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}\dot{\underset{\sim}{Z}}'\underset{\sim}{A}\dot{\underset{\sim}{Z}}\dot{\underset{\sim}{D}}_w\dot{\underset{\sim}{\Delta}}\dot{\underset{\sim}{D}}_w^{\frac{1}{2}}) \\
&= \text{tr}(\underset{\sim}{A}\underset{\sim}{V}_w)^2 + \text{tr}(\dot{\underset{\sim}{\Delta}}\dot{\underset{\sim}{D}}_w)^2 - 2\text{tr}(\underset{\sim}{A}\dot{\underset{\sim}{Z}}\dot{\underset{\sim}{D}}_w\dot{\underset{\sim}{\Delta}}\dot{\underset{\sim}{D}}_w\dot{\underset{\sim}{Z}}'), \quad \text{from (6.8)} \\
&= \text{tr}(\underset{\sim}{A}\underset{\sim}{V}_w)^2 + \sum_{i=0}^{c} \text{tr}[(p_i/q_i)^2 w_i^2 \underset{\sim}{I}_{q_i}] - 2\sum_{i=0}^{c} \text{tr}[\underset{\sim}{A}\underset{\sim}{Z}_i w_i^2(p_i/q_i)\underset{\sim}{Z}_i'], \text{ using (6.4) and (6.8)} \\
&= \text{tr}(\underset{\sim}{A}\underset{\sim}{V}_w)^2 + \sum_i p_i^2 w_i^2/q_i - 2\sum_i p_i w_i^2 \text{tr}(\underset{\sim}{A}\underset{\sim}{Z}_i\underset{\sim}{Z}_i')q_i \\
&= \text{tr}(\underset{\sim}{A}\underset{\sim}{V}_w)^2 - \sum_i p_i^2 w_i^2/q_i , \quad \text{after using (6.2).}
\end{aligned} \tag{6.9}$$

Since (6.9) involves $\underset{\sim}{A}$ only in the term $\text{tr}(\underset{\sim}{A}\underset{\sim}{V}_w)^2$, minimization of (6.9) with respect to elements of $\underset{\sim}{A}$ involves only the minimization of $\text{tr}(\underset{\sim}{A}\underset{\sim}{V}_w)^2$. This is to be done subject to $\underset{\sim}{A} = \underset{\sim}{A}'$, $\underset{\sim}{A}\underset{\sim}{X} = \underset{\sim}{0}$ and $\text{tr}(\underset{\sim}{A}\underset{\sim}{Z}_i\underset{\sim}{Z}_i') = p_i$, for $i = 0, 1, \cdots, c$.

6.4. THE ESTIMATORS

The MINQUE estimator of $\mathbf{p}'\boldsymbol{\sigma}^2$ is $\mathbf{y}'\mathbf{A}\mathbf{y}$ for $\mathbf{A}$ that minimizes $\operatorname{tr}(\mathbf{A}\mathbf{V}_w)^2$ subject to

$$\mathbf{A} = \mathbf{A}', \quad \mathbf{A}\mathbf{X} = \mathbf{0} \quad \text{and} \quad \operatorname{tr}(\mathbf{A}\mathbf{Z}_i\mathbf{Z}_i') = p_i \text{ for } i = 0, \cdots, c. \tag{6.10}$$

Use $\mathbf{V}_w = \mathbf{L}^{-1}\mathbf{L}^{-1'}$ similar to (1.16). Then

$$\operatorname{tr}(\mathbf{A}\mathbf{V}_w)^2 = \operatorname{tr}(\mathbf{A}\mathbf{L}^{-1}\mathbf{L}^{-1'})^2 = \operatorname{tr}(\mathbf{L}^{-1'}\mathbf{A}\mathbf{L}^{-1})^2 \tag{6.11}$$

and conditions (6.10) are equivalent to

$$\mathbf{L}^{-1'}\mathbf{A}\mathbf{L}^{-1} = (\mathbf{L}^{-1'}\mathbf{A}\mathbf{L}^{-1})', \ \mathbf{L}^{-1'}\mathbf{A}\mathbf{L}^{-1}\mathbf{L}\mathbf{X} = \mathbf{0} \text{ and } \operatorname{tr}[\mathbf{L}^{-1'}\mathbf{A}\mathbf{L}^{-1}(\mathbf{L}\mathbf{Z}_i)(\mathbf{L}\mathbf{Z}_i)'] = p_i . \tag{6.12}$$

Therefore, minimizing $\operatorname{tr}(\mathbf{A}\mathbf{V}_w)^2$ subject to (6.10) is equivalent to minimizing $\operatorname{tr}(\mathbf{L}^{-1'}\mathbf{A}\mathbf{L}^{-1})^2$ subject to (6.12); and this is identical to the minimization theorem of Section 2.10 with the substitutions

$$\mathbf{L}^{-1'}\mathbf{A}\mathbf{L}^{-1} \text{ for } \mathbf{Q}; \quad \mathbf{L}\mathbf{X} \text{ for } \mathbf{X}; \quad \mathbf{L}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{L}' \text{ for } \mathbf{W}_i; \quad \text{and} \quad p_i \text{ for } t_i . \tag{6.13}$$

Hence, from (2.95), the desired value of $\mathbf{A}$ is given by

$$\mathbf{L}'^{-1}\mathbf{A}\mathbf{L}^{-1} = \sum_i \lambda_i \mathbf{M}\mathbf{L}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{L}'\mathbf{M} ,$$

i.e.,

$$\mathbf{A} = \Sigma\lambda_i\mathbf{L}'\mathbf{M}\mathbf{L}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{L}'\mathbf{M}\mathbf{L}' ; \tag{6.14}$$

where, from (2.96),

$$\{\operatorname{tr}(\mathbf{M}\mathbf{L}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{L}'\mathbf{M}\mathbf{L}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{L}')\}\boldsymbol{\lambda} = \mathbf{p} \tag{6.15}$$

for, in this case,

$$\mathbf{M} = \mathbf{I} - \mathbf{L}\mathbf{X}(\mathbf{X}'\mathbf{L}'\mathbf{L}\mathbf{X})^{-}\mathbf{X}'\mathbf{L}' . \tag{6.16}$$

But in (6.14) and (6.15) $\underset{\sim}{M}$ occurs only in the form $\underset{\sim}{L}'\underset{\sim}{M}\underset{\sim}{L}$ which, from (6.16), is

$$\underset{\sim}{L}'\underset{\sim}{M}\underset{\sim}{L} = \underset{\sim}{L}'\underset{\sim}{L} - \underset{\sim}{L}'\underset{\sim}{L}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{L}'\underset{\sim}{L}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{L}'\underset{\sim}{L} = \underset{\sim}{V}_W^{-1} - \underset{\sim}{V}_W^{-1}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}_W^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}_W^{-1} = \underset{\sim}{P}_W ,$$

similar to (2.26). Hence in (6.14) the required $\underset{\sim}{A}$ is

$$\underset{\sim}{A} = \Sigma\lambda_i\underset{\sim}{P}_W\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}_W \tag{6.17}$$

with the λ_i's coming from (6.15) in the form

$$\{\mathrm{tr}(\underset{\sim}{P}_W\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}_W\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\}\underset{\sim}{\lambda} = \underset{\sim}{p} . \tag{6.18}$$

Thus the estimator of $\underset{\sim}{p}'\underset{\sim}{\sigma}^2$ is

$$\underset{\sim}{p}'\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{y}'\underset{\sim}{A}\underset{\sim}{y} = \Sigma\lambda_i\underset{\sim}{y}'\underset{\sim}{P}_W\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}_W\underset{\sim}{y} \tag{6.19}$$

with the λ_i's given by (6.18). Define the matrix in (6.18) as $\underset{\sim}{F}_W$,

$$\underset{\sim}{F}_W \equiv \{\mathrm{tr}(\underset{\sim}{P}_W\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}_W\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\} \quad \text{for } i, j = 0, 1, \cdots, c \tag{6.20}$$

so that (6.18) is

$$\underset{\sim}{F}_W\underset{\sim}{\lambda} = \underset{\sim}{p} . \tag{6.21}$$

And similarly define the vector of quadratic forms in (6.19) as $\underset{\sim}{u}_W$

$$\underset{\sim}{u}_W \equiv \{\underset{\sim}{y}'\underset{\sim}{P}_W\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}_W\underset{\sim}{y}\} \quad \text{for } i = 0, 1, \cdots, c \tag{6.22}$$

so that (6.19) is

$$\underset{\sim}{p}'\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{\lambda}'\underset{\sim}{u}_W .$$

Therefore, from (6.21)

$$\underset{\sim}{p}'\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{p}'\underset{\sim}{F}_W^{-1}\underset{\sim}{u}_W , \tag{6.23}$$

and on letting $\underset{\sim}{p}'$ be, in turn, the rows of the identity matrix, (6.23) is equivalent to

$$\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{F}_w^{-1}\underset{\sim}{u}_w = \{\mathrm{tr}(\underset{\sim}{P}_w\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}_w\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\}^{-1}\{\underset{\sim}{y}'\underset{\sim}{P}_w\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}_w\underset{\sim}{y}\} \tag{6.24}$$

which is the same as

$$\{\mathrm{tr}(\underset{\sim}{P}_w\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}_w\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\}\hat{\underset{\sim}{\sigma}}^2 = \{\underset{\sim}{y}'\underset{\sim}{P}_w\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{P}_w\underset{\sim}{y}\} \tag{6.25}$$

for $i, j = 0, 1, \cdots, c$. These are the MINQUE estimators of the variance components, using weights w_i in the norm, as in (6.8) and thence (6.9). Since these weights are pre-assigned numbers, $\underset{\sim}{V}_w$, $\underset{\sim}{P}_w$ and hence $\underset{\sim}{F}_w$ are matrices that can be calculated, and so the solutions (6.25) can also be calculated – provided the w_i's are such that $\underset{\sim}{F}_w^{-1}$ exists.

Two particular sets of values of the w_i are of interest. One is $w_i = 1$ for all i, $i = 0, 1, \cdots, c$, giving $\underset{\sim}{V}_w = \sum_{i=0}^{c} \underset{\sim}{Z}_i\underset{\sim}{Z}_i'$. Another is $w_0 = 1$, $w_i = 0$ for $i = 1, 2, \cdots, c$, giving

$$\underset{\sim}{V}_w = \underset{\sim}{I} \quad \text{and} \quad \underset{\sim}{P}_w = \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}' = \underset{\sim}{M} \text{ of } (2.17). \tag{6.26}$$

In this case equations (6.25) become

$$\hat{\underset{\sim}{\sigma}}^2 = \{\mathrm{tr}(\underset{\sim}{M}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{M}\underset{\sim}{Z}_j\underset{\sim}{Z}_j')\}^{-1}\{\underset{\sim}{y}'\underset{\sim}{M}\underset{\sim}{Z}_i\underset{\sim}{Z}_i'\underset{\sim}{M}\underset{\sim}{y}\} . \tag{6.27}$$

One might, perhaps, refer to these estimators as "original" MINQUE since they are akin to those suggested in Rao [1970], the first of his four papers on this topic. In point of fact, w_0 can be any value in this development (with $w_i = 0$ for $i \geq 1$) and (6.27) will still be the estimator of $\dot{\underset{\sim}{\sigma}}^2$.

6.5. ITERATIVE MINQUE

It is noticeable that the occurrence of σ_i^2 in $\underset{\sim}{V}$ is paralleled, in the MINQUE procedure, by the occurrence of w_i in $\underset{\sim}{V}_w$. Suppose then, after calculating a MINQUE estimate $\hat{\underset{\sim}{\sigma}}^2$ from (6.25), that we use the values therein as weights w_i and calculate

(6.25) again; and then repeat this iterative process until two successive values of $\hat{\boldsymbol{\sigma}}^2$ are equal, to some degree of approximation. The resulting $\hat{\boldsymbol{\sigma}}^2$ is called the iterative MINQUE estimator. It has been named I-MINQUE by Brown [1976], who also shows that MINQUE and I-MINQUE estimators are asymptotically normal.

I-MINQUE estimates can also be thought of in terms of using $\dot{\boldsymbol{\sigma}}^2$ for $\mathbf{w}$ in (6.25), so that $\mathbf{P}_w = \mathbf{P}$, and (6.25) becomes

$$\{\mathrm{tr}(\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j')\}\dot{\boldsymbol{\sigma}}^2 = \{\mathbf{y}'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{y}\} \quad \text{for } i, j = 0, 1, \cdots, c. \qquad (6.28)$$

Equations (6.28) are then solved iteratively for $\dot{\boldsymbol{\sigma}}^2$, the solutions being the I-MINQUE estimators. Because they are obtained iteratively they do not have the properties used in deriving equations (6.25); i.e., they are not necessarily unbiased or "best" in any sense.

6.6. COMPARISON OF MINQUE AND REML

The REML equations of (5.19) are

$$\{\mathrm{tr}(\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\mathbf{Z}_j\mathbf{Z}_j')\}\hat{\boldsymbol{\sigma}}^2 = \{\mathbf{y}'\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\mathbf{y}\} \quad \text{for } i, j = 0, 1, \cdots, c.$$

These are exactly the same as (6.28), except for notational differences; with both sets of equations, solutions are found by iteration. Hence we have the conclusion that

$$\text{I-MINQUE estimators} \equiv \text{REML estimators.} \qquad (6.29)$$

Derivation of I-MINQUE (REML) estimators begins with using some pre-assigned value $\dot{\boldsymbol{\sigma}}_0^2$ for $\dot{\boldsymbol{\sigma}}^2$ in $\mathbf{P}$ in (6.28). Denoting this value by $\mathbf{w}$ means that the first round iterate is the solution to (6.25), and therefore it is a MINQUE. Thus we have the conclusion:

$$\text{First iterate from REML} \equiv \text{a MINQUE.} \qquad (6.30)$$

6.7. MINIMUM VARIANCE (UNDER NORMALITY) - MIVQUE

Development of (6.9), the norm that is to be minimized in the MINQUE procedure, does not depend on normality. It can be used regardless of any distribution assumptions. In contrast, on assuming normality, the procedure for obtaining minimum variance, as distinct from minimum norm (quadratic unbiased translation invariant) estimators is easily derived. This is so because, under normality, the variance of $\mathbf{y}'\mathbf{A}\mathbf{y}$ is

$$v(\mathbf{y}'\mathbf{A}\mathbf{y}) = 2\mathrm{tr}(\mathbf{A}\mathbf{V})^2 + 4\boldsymbol{\alpha}'\mathbf{X}'\mathbf{A}\mathbf{V}\mathbf{A}\mathbf{X}\boldsymbol{\alpha} = 2\mathrm{tr}(\mathbf{A}\mathbf{V})^2$$

after utilizing $\mathbf{A}\mathbf{X} = \mathbf{0}$ arising from translation invariance. Therefore, if the weights w_i used in MINQUE are deemed to be <u>a priori</u> values of the σ_i^2, then in order to obtain a minimum variance estimator $\mathbf{y}'\mathbf{A}\mathbf{y}$ of $\mathbf{p}'\boldsymbol{\sigma}^2$ that is translation invariant and unbiased, we need to minimize $\mathrm{tr}(\mathbf{A}\mathbf{V})^2$ subject to $\mathbf{A} = \mathbf{A}'$, $\mathbf{A}\mathbf{X} = \mathbf{0}$ and $\mathrm{tr}(\mathbf{A}\mathbf{Z}_i\mathbf{Z}_i') = p_i$ for $i = 0, 1, \cdots, c$, the latter coming from (6.1) and (6.2), respectively. But this restricted minimization of $\mathrm{tr}(\mathbf{A}\mathbf{V})^2$ is precisely the same as that of $\mathrm{tr}(\mathbf{A}\mathbf{V}_w)^2$ described following (6.9). Therefore

$$\text{MIVQUE (under normality)} \equiv \text{I-MINQUE} \equiv \text{REML} . \tag{6.31}$$

REFERENCES

Anderson, R. D. [1977]. Personal communication.

Anderson, R. L. and Bancroft, T. A. [1952]. Statistical Theory in Research. McGraw Hill, New York.

Anderson, T. W. [1969]. Statistical inference for covariance matrices with linear structure, in Multivariate Analysis, 55-66, Ed. P. R. Krishnaiah, Academic Press, New York.

Anderson, T. W. [1970]. Estimation of covariance matrices which are linear combinations or whose inverses are linear combinations of given matrices, 1-24, in Bose, R. C., Chakravarti, I. M., Mahalanobis, P. C., Rao, C. R., and Smith, K. J. C., Eds., Essays in Probability and Statistics, University of North Carolina Press, Raleigh, North Carolina.

Brown, K. G. [1976]. Asymptotic behavior of MINQUE-type estimators of variance components. Annals of Statistics 4:746-754.

Corbeil, R. R. and Searle, S. R. [1976a]. Restricted maximum likelihood (REML) estimation of variance components in the mixed model. Technometrics 18:31-38.

Corbeil, R. R. and Searle, S. R. [1976b]. A comparison of variance components estimators. Biometrics 32:779-791.

Daniels, H. E. [1939]. The estimation of components of variance. J. Roy. Stat. Soc. (Supp.) 6:186-197.

Graybill, F. A. and Hultquist, R. A. [1961]. Theorems concerning Eisenhart's Model II. Ann. Math. Stat. 32:261-269.

Graybill, F. A. and Wortham, A. W. [1956]. A note on uniformly best unbiased estimators for variance components. J. Am. Stat. Assoc. 51:266-268.

Hartley, H. O. and Rao, J. N. K. [1967]. Maximum likelihood estimation for the mixed analysis of variance model. Biometrika 54:93-108.

Harville, D. A. [1974]. Bayesian inference for variance components using only error contrasts. Biometrika 61:383-385.

Harville, D. A. [1975, 1977]. Maximum likelihood approaches to variance component estimation and to related problems. Technical Rpt. No. 75-0175, Aerospace Res. Lab, Wright-Patterson AFB, Ohio, 1975. Revised and up-dated, J. Am. Stat. Assoc. 72:320-340.

Harville, D. A. [1976]. Extension of the Gauss-Markov theorem to include the estimation of random effects. Ann. of Stat. 4:384-395.

Hemmerle, W. J. and Hartley, H. O. [1973]. Computing maximum likelihood estimates for the mixed A.O.V. model using the W-transformation. Technometrics 15:819-832.

Henderson, C. R. [1973]. Maximum likelihood estimation of variance components. Unpublished manuscript, Animal Science Department, Cornell University.

Henderson, C. R., Kempthorne, O., Searle, S. R., and von Krosig, C. M. [1959]. The estimation of environmental and genetic trends from records subject to culling. Biometrics 15:192-218.

Jennrich, R. I. and Sampson, P. F. [1976]. Newton-Raphson and related algorithms for maximum likelihood estimation of variance components. Technometrics 18: 11-18.

Khatri, C. G. [1966]. A note on a manova model applied to problems in growth curve. Ann. Inst. Stat. Math. 18:75-86.

LaMotte, L. R. [1973]. Quadratic estimation of variance components. Biometrics 29:311-330.

MacRae, E. C. [1974]. Matrix derivatives with an application to an adaptive linear decision problem. Ann. of Stat. 2:337-346.

Marsaglia, G. and Styan, G. P. H. [1974a]. Equalities and inequalities for ranks of matrices. Linear and Multilinear Algebra 2:269-292.

Marsaglia, G. and Styan, G. P. H. [1974b]. Rank conditions for generalized inverses of partitioned matrices. Sankhyā 36:437-442.

Miller, J. J. [1973]. Asymptotic properties and computation of maximum likelihood estimates in the mixed model of the analysis of variance. Tech. Report 12, Dept. of Statistics, Stanford University, Stanford, California.

Neudecker, H. [1969]. Some theorems on matrix derivatives in multivariate analysis. J. Am. Stat. Assoc. 62:607-625.

Patterson, H. D. and Thompson, R. [1971]. Recovery of interblock information when block sizes are unequal. Biometrika 58:545-554.

Patterson, H. D. and Thompson, R. [1974]. Maximum likelihood estimation of variance components. Proceedings of the 8th International Biometric Conference, 197-207.

Rao, C. R. [1970]. Estimation of heteroscedastic variances in a linear model. J. Am. Stat. Assoc. 65:165-172.

Rao, C. R. [1971a]. Estimation of variance and covariance components - MINQUE theory. J. of Multivariate Analysis 1:257-275.

Rao, C. R. [1971b]. Minimum variance quadratic unbiased estimation of variance components. J. of Multivariate Analysis 1:445-456.

Rao, C. R. [1972]. Estimation of variance and covariance components in linear models. J. Am. Stat. Assoc. 67:112-115.

Roth, W. E. [1934]. On direct product matrices. Bull. Amer. Math. Soc. 40:461-468.

Schaeffer, L. R. [1976].

Searle, S. R. [1966]. Matrix Algebra for the Biological Sciences. Wiley, New York.

Searle, S. R. [1970]. Large sample variances of maximum likelihood estimators of variance components using unbalanced data. Biometrics 26:505-524.

Searle, S. R. [1971a]. Topics in variance component estimation. Biometrics 27:1-76.

Searle, S. R. [1971b]. Linear Models. Wiley, New York.

Searle, S. R. [1978]. A univariate formulation of the multivariate linear model. Survey Sampling and Applied Statistics: of H. O. Hartley. Ed., H. A. David.

Thompson, W. A. [1962]. The problem of negative estimates of variance components. Ann. Math. Stat. 33:273-289.

Townsend, E. C. [1968]. Unbiased estimators of variance components in simple unbalanced designs. Ph.D. Thesis, Cornell University, Ithaca, New York.

Townsend, E. C. and Searle, S. R. [1971]. Best quadratic unbiased estimation of variance components from unbalanced data in the 1-way classification. Biometrics 27:643-657.

Tracy, D. S. and Dwyer, P. S. [1969]. Multivariate maxima and minima with matrix derivatives. J. Am. Stat. Assoc. 64:1576-1594.

Winsor, C. P. and Clarke, G. L. [1940]. A statistical study of variation in the catch of plankton nets. Sears Foundation J. of Marine Res. 3:1-34.

Zacks, S. [1971]. The Theory of Statistical Inference. Wiley, New York.